THE ISOLATION AND CHARACTERISATION OF NUCLEAR-ENCODED LIGHT-REGULATED GENES FROM PISUM SATIVUM AND THEIR EXPRESSION IN TRANSGENIC PLANTS OF NICOTIANA TABACUM.

Philip Mark Gilmartin BSc (Hons) Leeds

A thesis presented for the degree of
Doctor of Philosophy

Department of Biological Sciences
University of Warwick

November 1986

to Mum and Dad

CONTENTS

CONTENTS

SECTION III - RESULTS AND DISCUSSION

LIST OF FIGURES

LIST OF TABLES

LIST OF COLOUR PLATES

ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor R. John Ellis F.R.S. for his guidance and encouragement during the course of this work. I am greatly indebted to Dr. T. F. Gallagher not only for providing the lambda genomic clones upon which this work is based, but also for giving me so much practical guidance and teaching me so much.

I am grateful to Dr. M.W. Bevan for providing the Agrobacterium binary vectors along with the necessary practical information on how to use them, and to Miss S.E. Mason for showing me the relevant plant tissue culture techniques during my visit to the Plant Breeding Institute, Cambridge.

I would also like to thank the many members of the Department of Biological Sciences at Warwick who have provided me with help and advice over the last three years.

I am grateful to the SERC for providing me with funding in the form of an SERC Biotechnology Directorate studentship.

Finally I thank my parents for all their encouragement and support; without them, none of this would have been possible.

DECLARATION

The work presented in this thesis is entirely my own and was conducted under the supervision of Professor R.J. Ellis F.R.S. None of the results have been used in any previous application for a degree. All published work has been acknowledged by reference.

21.11.86

ABBREVIATIONS

A	absorbance
AB genes	genes encoding the major chlorophyll a/b binding protein
AMPS	ammonium persulphate
A.tumefaciens	Agrobacterium tumefaciens
BA	6-benzylaminopurine
Bis	N,N'-methylenebisacrylamide
BME	2-mercaptoethanol
BPB	bromophenol blue
BSA	bovine serum albumin
CIP	calf intestinal phosphatase
DNA	deoxyribonucleic acid
DTT	dithiothreitol
DW	distilled water
dATP	2'-deoxyadenosine 5'-triphosphate
[α-^{35}S]αSdATP	2'-deoxyadenosine 5'-[α-^{35}S]thiotriphosphate
dCTP	2'-deoxycytosine 5'-triphosphate
ddATP	2',3'-dideoxyadenosine 5'-triphosphate
ddCTP	2',3'-dideoxycytidine 5'-triphosphate
ddGTP	2',3'-dideoxyguanosine 5'-triphosphate
ddTTP	2',3'-dideoxythymidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
[α-^{32}P]dGTP	2'-deoxyguanosine 5'-[α-^{32}P]triphosphate
dTTP	2'-deoxythymidine 5'-triphosphate
E.coli	Escherichia coli
EDTA	ethylenediaminetetra-acetic acid
HCC	hexamine cobalt chloride
Hepes	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid

IPTG	isopropyl-B-D-thio-galactoside
h	hours
LB	Luria-Bertani medium
LHCP	major chlorophyll a/b binding protein
LMP	low melting point
LS	large subunit of rubisco
Mes	2-(N-morpholino)ethanesulfonic acid
min	minute
Mops	3-(N-morpholino)propanesulfonic acid
mRNA	messenger ribonucleic acid
MS	Murashige & Skoog
NAA	α-naphthalene acetic acid
NB	nutrient broth
N.tabacum	Nicotiana tabacum
PAGEB	polyacrylamide gel elution buffer
PB	phage buffer
PEG	polyethylene glycol 6000
Pipes	piperazine-N,N'-bis(2-ethanesulfonic acid)
pfu	plaque forming unit
Poly (A)	polyadenylic acid
P.sativum	Pisum sativum
pvp40	polyvinyl pyrolidone 40
RF DNA	replicative form deoxyribonucleic acid
RNA	ribonucleic acid
rubisco	ribulose bisphosphate carboxylase/oxygenase
SB	super broth
SDW	sterile distilled water
SDS	sodium dodecyl sulphate
SS	small subunit of rubisco
SSC	standard saline citrate

ssDNA	single-stranded deoxyribonucleic acid
TE	Tris-EDTA buffer
TEMED	N,N,N',N'-tetramethylethylenediamine
Tfb	transformation buffer
TNS	tri-iso-propylnaphthalenesulfonic acid (Na salt)
Tris	2-amino-2(hydroxymethyl)-1,3-propanediol
tRNA	transfer ribonucleic acid
u	unit
u.v.	ultra-violet
X-Gal	5-bromo-4-chloro-3-indolyl-ß-galactoside

SUMMARY

Genomic clones encoding a gene for the small subunit of ribulose-1,5-bisphosphate carboxylase oxygenase (SS47) and a gene encoding the major chlorophyll *a*/*b* binding protein (AB13) from *Pisum sativum* have been isolated and characterised. The genomic clone SS47 is shown to be the genomic equivalent of the cDNA clone SSU60 [Bedbrook, Smith & Ellis (1980) Nature 287, 692-697.] by hybridisation studies and sequence analysis. The complete sequence of SS47 and the partial sequence of AB13 are presented.

The gene SS47 accounts for 50% of all small subunit transcripts in 9-day-old light-grown pea leaves. The degree of expression of the other members of the small subunit multi-gene family is also discussed. The gene AB13, or a highly homologous one, is shown to be expressed in 9-day-old light-grown pea leaf tissue.

The characterised genes have been transferred, via an *Agrobacterium* Ti-plasmid binary vector to cells of *Nicotiana tabacum*. Transformed cells have been regenerated into whole plants. The degree of expression of the gene SS47 in transgenic tobacco plants is shown to be variable between plants, and ranges from zero up to levels comparable to its expression in pea plants. In the one transgenic plant containing AB13 that was screened, no AB13 transcripts were detected.

Studies on the transmission of the kanamycin-resistance marker to the progeny of transformed plants have shown that in all but one of these plants, segregation of this marker is in a typical Mendelian manner; a possible explanation is proposed and discussed for the non-Mendelian segregation of the kanamycin-resistance marker in the progeny of one transformed plant.

SECTION I

LITERATURE REVIEW

1. INTRODUCTION

The growth and development of plants involves the coordinated expression of three distinct genetic systems located within the nucleus, chloroplast and mitochondrion. The chloroplast and mitochondrial genomes do not however encode all of the proteins found within these organelles; the nuclear genome contributes most of these.

In recent years, with the advances in molecular biological techniques, plant genomes have become amenable to investigation and characterisation. The isolation of cDNA clones for nuclear genes encoding chloroplast proteins has not only provided valuable information on the expression, regulation and organisation of these genes, but has enabled the isolation of the nuclear genes themselves.

In order to gain an understanding of nuclear gene expression and regulation, a means of introducing characterized and modified genes back into plant cells is required. The recent development of plant transformation vectors has provided the necessary technology to do this. These vectors are the most powerful tools currently available to the plant molecular biologist and open the door, not only to the analysis of sequences involved in gene regulation, but also to the genetic modification of plants.

This thesis presents the results of experiments aimed at understanding the regulation of nuclear genes encoding two chloroplast-associated proteins in both normal and transgenic plants. These proteins are the small subunit of the enzyme ribulose-1,5-bisphosphate

carboxylase-oxygenase (Rubisco, EC4.1.1.39), and the major chlorophyll a/b binding protein (LHCP).

In the following literature review I will discuss those aspects of plant biochemistry and molecular biology that led up to, and have emerged during, the work presented here on these two proteins.

The final section of this review will discuss the aims of the present work. These aims will be considered in relation to the information that was available at the onset of this study since much of the information discussed in this review has been published during the course of the work presented in this thesis.

2. RIBULOSE BISPHOSPHATE CARBOXYLASE-OXYGENASE

2.A Structure and activity

Ribulose-1,5-bisphosphate carboxylase-oxygenase (Rubisco) is the enzyme which catalyses the first reaction in both photosynthetic carbon assimilation and photorespiration (Miziorko and Lorimer, 1983). This enzyme is not only the most abundant soluble protein in plant cells, but is perhaps the most abundant enzyme in the world (Ellis, 1979). The enzyme is found in the stroma of the chloroplast where both the carboxylation and oxygenation reactions occur. The carboxylation reaction catalysed by Rubisco is summarised as follows:-

$$CO_2 + \text{ribulose-1,5-bisphosphate} + H_2O \xrightarrow{Mg^{2+}} 2\ \text{3-phosphoglycerate}$$

The 3-phosphoglycerate is subsequently phosphorylated with ATP to produce 3-phosphoglyceryl phosphate. This product is reduced by NADPH to glyceraldehyde 3-phosphate. The glyceraldehyde 3-phosphate equilibrates with dihydroxyacetone phosphate to form a stromal pool of triose phosphates. From this pool, hexose sugars are generated and ribulose-1,5-bisphosphate is regenerated.

The second reaction catalysed by Rubisco, is an oxygenation reaction and can be summarised as follows:-

$$O_2 + \text{ribulose-1,5-bisphosphate} \xrightarrow{Mg^{2+}} \text{3-phosphoglycerate} + \text{2-phosphoglycollate}$$

The 2-phosphoglycollate is then de-phosphorylated to generate glycollate. The glycollate is subsequently metabolised by both the peroxisomes and mitochondria, leading to the evolution of CO_2. Both CO_2 and O_2 compete at the active site of Rubisco for ribulose-1,5-bisphosphate. Photorespiration is regarded as being not only wasteful of energy but also wasteful of fixed carbon (for review see Miziorko & Lorimer, 1983).

The enzyme Rubisco contains two types of subunit, termed large subunit (LS) and small subunit (SS) (Miziorko & Lorimer, 1983). The large subunit is encoded by the chloroplast genome (Ellis, 1981) and may be synthesised as a precursor protein (Langridge, 1981). The mature LS protein has a molecular weight of approximately 56,000 (Miziorko & Lorimer, 1983). The small subunit is encoded within the nucleus and is synthesised on free cytoplasmic ribosomes as a precursor protein containing a 40 to 60 aminoacyl transit sequence at the amino terminus (for review see Miziorko & Lorimer, 1983). The SS precursor is transported post-translationally to the chloroplast where the transit peptide is cleaved by a specific endo-protease to yield the mature SS polypeptide of approximately 14 kDa (Robinson and Ellis, 1984).

In higher plants, the enzyme Rubisco consists of eight small and eight large subunits (Miziorko and Lorimer, 1983). The assembly of the holoenzyme is thought to involve a third protein, of nuclear origin, termed the LS-binding protein. This binding protein consists of six of each of two subunit types, α and β (Musgrove and Ellis, 1986).

The LS contains the active site of the enzyme whilst

no catalytic function is known for the SS. The SS is however essential for enzyme activity and is thought to play a role in maintaining the protein in an active conformation (Miziorko & Lorimer, 1983).

2.B Gene organisation

In this Section, the organisation of the genes encoding both the large subunit (LS) and the small subunit (SS) of Rubisco from higher plants will be discussed. The location and copy-number of the LS genes will be considered first, followed by the location and copy-number of the SS genes. A detailed survey of the SS genes from various monocotyledonous and dicotyledonous plants will be made. The sequences involved in the regulation of the SS genes will be discussed only briefly here and will be considered in greater detail in Section I.2.C.

The gene encoding the LS of Rubisco is present as a single copy per chloroplast DNA molecule. The copy-number of chloroplast DNA molecules per chloroplast varies between 10 and 800, depending on the species. The copy-number of LS genes per chlorplast may therefore be as high as 800. The number of chloroplasts per plant cell is variable and dependant on the cell type. In leaf tissue, there may be as many as 200 chloroplasts per cell. The resulting copy-number of LS genes may therefore be several thousand per cell (Miziorko and Lorimer, 1983).

The genes encoding the SS of Rubisco are located within the nuclear genome (Kawashima and Wildman, 1972) as a small multi-gene family (Berry-Lowe et al., 1982; Broglie et al., 1983; Cashmore et al., 1983; Coruzzi et al., 1983;

Dunsmuir et al., 1983a; Wimpee et al., 1983; Coruzzi et al., 1984; Pichersky et al., 1986). The number of copies of the SS genes in the multi-gene family varies between species. Tomato (Lycopersicon esculentum) contains only three SS genes per haploid genome (Pichersky et al., 1986) whilst duckweed (Lemna gibba) contains approximately thirteen (Wimpee et al., 1983). Other plant species contain between three and thirteen SS genes per haploid genome. Numerous SS genes have been cloned from several plant species: these are summarised in Table 1. The total number of SS genes isolated from higher plants is twenty one genomic sequences and twelve cDNA sequences. The genomic sequences corresponding to nine of the cDNA sequences are not known. There are therefore thirty different characterised SS genes from eight different plant species. The abundance of data concerning SS genes probably makes them the best characterised of all plant genes.

Several authors have shown, by the presence of more than one SS gene on DNA fragments cloned into phage λ, that different members of the SS multi-gene family are physically linked (Cashmore, 1983; Dean et al., 1985b; Tobin et al., 1985). The chromosomal location of the SS genes in pea has been investigated by Polans et al. (1985). These authors have shown that the SS multi-gene family segregates as a single Mendelian unit on chromosome 5. This genetic observation is consistent with the data obtained from the molecular studies on the SS multi-gene family.

The SS genes from both monocotyledonous and dicotyledonous plant species contain structural similarities. The SS genes from all dicotyledonous plants

Table 1. Characterised Rubisco SS sequences

The plant species, the number of cDNA and genomic clones that have been isolated from that species, and the relevant references are shown in the table below.

Species	cDNA Clones	Reference
Lemna gibba	one	Stiekema et al., 1983a
Lycopersicon esculentum	one	Pichersky et al., 1986
Nicotiana sylvestris	one	Pinck et al., 1983
Petunia hybrida	three	Dunsmuir et al., 1983a
Pisum sativum	three	Bedbrook et al., 1980
		Broglie et al., 1981
		Coruzzi et al., 1983
Triticum aestivum	four	Broglie et al., 1983
		Smith et al., 1983

Species	Genomic Clones	Reference
Glycine max	one	Berry-Lowe et al., 1982
Lemna gibba	five	Wimpee et al., 1983
Lycopersicon esculentum	two	Pichersky et al., 1986
Nicotiana tabacum	one	Mazur and Chui, 1985
Petunia hybrida	seven	Dean et al., 1985b
Pisum sativum	four	Cashmore, 1983
		Coruzzi et al., 1984
		Timko et al., 1985a
		This work
Triticum aestivum	one	Broglie et al., 1983

studied to date contain at least two introns which occur at precisely the same positions in all species. The first intron is located 6 bp 3' to the transit peptide cleavage point, between the 6th and 7th nucleotide of the mature SS polypeptide. This intron is of variable size between different species and between different SS genes within species. In petunia, the size of this intron ranges from 106 bp to 970 bp (Dean et al., 1985a). In pea the size range is between 75 bp and 471 bp (Coruzzi et al., 1983; Timko et al., 1985; this work). The size of this intron is also variable in other species (Pichersky et al., 1986). The second intron is situated 135 bp 3' to the first intron. In petunia, the size of this second intron varies between 121 bp and 310 bp, whilst in pea it is more highly conserved in both size (86 bp to 87 bp) and sequence between different members of the gene family (Coruzzi et al., 1983; Timko et al., 1985; this work). Two SS genes have been shown to contain a third intron. These genes have been isolated from petunia (Dean et al., 1985a) and tobacco (Mazur and Chui, 1985). The position of the third intron is identical in both these species; namely 53 bp 3' to the second intron. This position also coincides with the site of insertion of 12 amino acids into higher plant SS genes as compared to those of the blue-green algae Anacystis (Shinozaki and Sugiura, 1983) and Anabaena (Nierzwicki-Bauer et al., 1984). The significance of the size and number of introns in these genes is not known. It has however been noted that the petunia SS gene containing three introns is the most highly expressed of all petunia SS genes (Dean et al., 1985a); the pea SS gene containing the longest first intron is also the

most highly expressed SS gene in pea (this work).

In contrast to the arrangement of introns in the SS genes of dicotyledonous plants, the SS genomic sequences of monocotyledonous plants contain only one intron. In wheat, this intron is located in the same position as the first intron of dicotyledonous plants (Broglie et al., 1983). The position of the single intron in Lemna is at the same location of the second intron found in all SS genes from dicotyledonous plants (Tobin et al., 1985). The implication from the observed variation in intron number both between SS genes from different species and between SS genes from the same species is that the gene duplication events that have led to the formation of the SS multi-gene families have been accompanied by either the loss, or possible gain, of intervening sequences.

The 3'-untranslated region of SS genes from all species is between 150 bp and 200 bp in length; different members of the multi-gene family can be distinguished by their sequence variation within this region (Dean et al., 1985c). The sequence variation within the 3'-untranslated region can be exploited to identify the genomic sequences corresponding to various cDNA probes (Coruzzi et al., 1984; this work). The sequence differences within this region can also be used to classify the SS genes into sub-families (Wimpee et al., 1983; Dean et al., 1985b; this work). It is interesting to note that the SS 3'-flanking sequences do not generally show sequence conservation between species. However, of the three SS genes isolated from tomato (Pichersky et al., 1986), one sequence shows greater sequence homology within its 3'-untranslated region to the

corresponding region from a tobacco SS gene than it does to the 3'-untranslated region of the other tomato genes (Pichersky et al., 1986). The evolutionary implication of this observation is that the SS genes of tomato diverged from each other earlier than the division of Lycopersicon esculentum and Nicotiana tabacum.

The use of chimaeric genes and Agrobacterium-mediated plant transformation (see Section I.4) has enabled the identification of sequences located 5' to the coding region of the SS genes from pea that are responsible for light-regulation and organ-specific expression. It has been shown that 352 bp of sequence 5' to the transcriptional start site are sufficient for the light-regulated expression of a pea SS gene in leaf tissue of transgenic tobacco plants (Nagy et al., 1985). Similar results have recently been obtained indicating that 410 bp directly 5' to the transcriptional start site of a different pea SS gene are sufficient for its correct regulation and expression in transgenic petunia plants (Fluhr and Chua, 1986). Within this upstream region from pea SS genes, an enhancer-like element has been identified that confers high level expression and photo-regulation upon the these genes (Timko et al., 1985b; Fluhr et al., 1986). Homology within this region to the viral enhancer core sequence has been noted both in pea (Fluhr and Chua, 1986) and petunia (Dean et al., 1985a). These observations will be discussed in greater detail in relation to the regulation of SS genes in Section I.2.C.

Having considered the similarities between different SS genes at the DNA sequence level, a review of their

conserved protein sequences will now be made. The aminoacyl sequence of the mature SS polypeptides from pea are all identical with the exception of SSU1 (Bedbrook et al., 1980) which differs by eight amino acids from PS2.4 (Coruzzi et al., 1984), SS3.6 (Cashmore, 1983), SS8.0 (Timko et al., 1985) and SS47 (this work). The transit peptides encoded by PS2.4 (Coruzzi et al., 1984) and SS3.6 (Cashmore, 1983) are identical. Compared to this sequence, the transit peptide encoded by SS8.0 contains one amino acid substitution and that encoded by SS47 contains two amino acid substitutions (see Section III.3.D). The SS genes of petunia all encode identical mature polypeptides (Dean et al., 1985c). The mature peptides encoded by SS genes shows greater sequence conservation, both within and between species, than do the transit peptides. A comparison of the SS transit peptides from all species indicates a short region of homology surounding the cleavage point (Broglie et al., 1983; Mazur and Chui, 1985). Several other amino acids appear to be conserved in the SS transit peptides, but no further major homology blocks exist.

The mature SS peptides show several regions of homology between species. These regions are also conserved between SS polypeptides from monocotyledonous and dicotyledonous plants. The major conserved region consists of 16 aminoacyl residues, identical in all higher SS polypeptides examined to date. This highly conserved region extends from Tyr^{61} to Gly^{76} in all species; 15 of these 16 amino acids are also conserved in the SS polypeptide from Euglena gracilis (Sailland et al., 1986). It is also interesting to note that 9 out of 10 of the amino acids

between Tyr^{66} and Phe^{75} are also conserved in SS polypeptides of the blue-green alga Anacystis nidulans (Shinozaki and Sugiura, 1983). The first five amino acids of the block of sixteen are, however, absent in blue-green algal SS polypeptides. Higher plant SS genes encode an extra twelve amino acids at this point as compared to blue-green algal genes; the last five of these twelve amino acids form the start of the sixteen amino acid conserved region.

A second region, showing slightly less sequence conservation, exists in higher plant SS polypeptides from amino acid 10 to amino acid 22. This region is also conserved between the Euglena, and the blue-green algal SS polypeptides. The significance of these highly conserved regions is as yet unclear. They may however be involved in the binding of the LS and the maintenance of the quaternary structure of the enzyme.

The two SS multi-gene families that have been studied in the greatest detail are those of petunia and pea. The petunia family contains eight genes; genomic sequences encoding all these genes have been isolated (Dean et al., 1985b). The eight SS family members fall into three sub-families according to homologies within their 3'-untranslated regions. The pea SS gene family contains at least five members, four of which have been isolated as genomic sequences. These sequences are: SS3.6 (Cashmore, 1983); PS2.4, also referred to as SS4.0 and rbcS-E9 (Coruzzi et al., 1984); SS8.0 (Timko et al., 1985) and SS47 (this work). Three cDNA sequences of pea SS genes have been isolated. These sequences are: SSU1 (Bedbrook et al., 1980) and SSU60 (Bedbrook et al., 1980), this sequence has since

been shown to differ from that of SSU1 in its 3'-untranslated region (S.M.Smith, unpublished results). The third pea SS cDNA sequence is SS15 (Broglie et al., 1981; Coruzzi et al., 1983). SS15 is the genomic equivalent of PS2.4 (Coruzzi et al., 1984) and SSU60 is the genomic equivalent of SS47 (this work). Two further SS genomic sequences have been isolated. These sequences are referred to as rbcS-3A and rbcS-3C (Fluhr and Chua, 1986). The complete nucleotide sequences of these genes have not yet been published.

2.C Gene expression

The accumulation of the enzyme Rubisco is stimulated by light (Highfield and Ellis, 1978; Ellis, 1981). In some plant species, for example pea and Lemna (Tobin and Silverthorne, 1985), the amounts of Rubisco are significantly lower in etiolated as opposed to light-grown plants. In cereals, however, Rubisco is not only the most abundant soluble protein in light-grown tissue, but also in etiolated tisue (Ellis, 1981); the amount of protein is however increased in response to light. In pea, the amounts of LS mRNA are enhanced in light-grown as compared to etiolated seedlings (Smith and Ellis, 1981). In maize, however, there is little difference between the amounts of LS mRNA in dark-grown and light-grown plants (Tobin and Silverthorne, 1985). The presence of two different-sized LS transcripts in maize, that alter in relative abundance during light-induced plastid development, may indicate an alternative mechanism for the control of LS polypeptide accumulation.

The first observation that changes in SS polypeptide amount are correlated to changes in the amount of translatable mRNA came from the *in vitro* translation of poly (A) RNA from light- and dark-grown plants (Tobin, 1978; Tobin and Suttie, 1980; Sasaki *et al.*, 1981; Smith and Ellis, 1981). The isolation of cDNA clones encoding the SS of Rubisco enabled the conclusion that the changes in the amount of translatable mRNA are correlated to increased amounts of mRNA, and not the conversion of untranslatable SS mRNA into a translatable form (Bedbrook *et al.*, 1980; Stiekema *et al.*, 1983b). The regulation of SS mRNA has been shown directly to be at the level of transcription by the use of run-off transcription in isolated nuclei (Gallagher and Ellis, 1982; Silverthorne and Tobin, 1984; Gallagher *et al.*, 1985).

The transcriptional regulation of SS genes from *Lemna* (Tobin, 1981), soybean (see Tobin and Silverthorne, 1985) and pea (Bennett *et al.*, 1984; Fluhr and Chua, 1986) is mediated by phytochrome, especially in etiolated tissue.

Phytochrome is a photoreceptor that controls plant development in response to light. The molecule has two photo-interconvertible forms, a red-absorbing form, P_R, that is biologically inactive and a far-red-absorbing form, P_{FR}, that is biologically active. In its far-red form phytochrome can modulate the expression of many light-regulated plant genes; classical phytochrome-modulated gene regulation shows red-light inducibility that is reversible by far-red-light (Pratt, 1982)

The SS genes from different plant species exhibit different temporal responses to light. In *Lemna* for example,

the increase of SS mRNA after light-induction is rapid (Tobin and Suttie, 1980) whilst in pea, there is a lag of approximately 24 hours (Jenkins et al., 1983). The rapid decline in SS transcription in pea on transfer of the plant to the dark contrasts with the slower decline in the level of transcription in similarly treated Lemna and soybean plants (Tobin and Silverthorne, 1985). The amount of SS mRNA is not only increased by light, but is induced by cytokinin (Funckes-Shippy and Levine, 1985).

The presence of Rubisco and its RNA in leaves is associated only with chloroplast-containing tissue. It is possible therefore that SS gene transcription may not only be influenced by phytochrome and cytokinin levels, but also by the development of chloroplasts (Simpson et al , 1986b).

It has been observed that in mature pea leaves, continuous illumination with red light alone cannot induce an increase in SS mRNA amounts, whilst a pulse of blue light can cause an increase in the SS mRNA amounts in mature plants (Kaufman et al., 1985). Fluhr and Chua (1986) have demonstrated that in etiolated primary leaves, changes in SS transcript levels are mediated by phytochrome. In mature green leaves, the level of SS transcripts are modulated by a blue light receptor acting in concert with phytochrome.

Rubisco is not present in all tissues of light-grown plants. Its presence is closely linked with green chloroplast-containing tissue (Simpson et al., 1986b). For this reason, the expression of SS genes appears to be organ-specific (Coruzzi et al., 1984); it is in fact not only organ-specific but cell-type specific (Simpson et al., 1986b). These workers have shown, by the transfer of

chimaeric SS-neomycin phosphotransferase (SS-NPT II) genes into tobacco plants, that the SS 5'-flanking region directs the expression of this chimaeric gene only in the chloroplast-containing tissue of leaves.

The ability to distinguish between different members of the SS multi-gene family, by sequence divergence between their 3'-untranslated regions, has enabled the relative degree of their expression to be investigated. Dean et al. (1985a) have shown that of the eight SS genes present in petunia, two show no detectable expression, five give rise to between 2 and 23% of the total SS transcripts in leaf tissue and one accounts for 47% of all SS transcripts in leaf tissue. The reason for such a range in the relative transcript amounts for different SS genes is not fully understood, but may involve differences either in the transcription rates of different genes, or in the stability of their transcripts. It has been shown, however, that linkage of the SS genes shows no correlation with the degree to which they are expressed (Dean et al., 1985a).

The work of Coruzzi et al. (1984) demonstrated that one pea SS gene is responsible for approximately 30 to 35% of total SS transcripts in green leaves. Two other pea SS genes are known to be transcriptionally active from the isolation of their cDNA clones. These genes are the genomic equivalents of SSU1 and SSU60 (Bedbrook et al., 1980). The genomic equivalent of SSU1 has not been isolated. The genomic equivalent of SSU60 is SS47 which accounts for over 50% of SS transcripts (see Section III.4). The other pea SS genes (SS3.6, SS8.0 and SSU1) account for between less than 5% and 10% of the total SS transcripts in pea leaves (see

Section III.4). Different pea SS genes therefore contribute in varying degrees to the total amount of SS transcripts. There appears to be no differential expression of petunia SS genes between different organs (Dean et al., 1985a), whilst in pea, some members of the gene family are under-represented in certain organ types (Coruzzi et al., 1984; Simpson et al., 1986b).

Recent work from several groups has demonstrated that the light-regulation and organ-specificity of SS genes from several plant species can be reproduced in transgenic plants of different species (Broglie et al., 1984; Herrera-Estrella et al., 1984; Nagy et al., 1985; Timko et al., 1985; Fluhr and Chua, 1986). These observations indicate that the DNA regions involved in the light-induced and organ-specific expression of the SS genes are highly conserved between different plant species.

Much work has gone into identifying the regions involved in the expression of the SS genes. The use of the 5'-untranslated regions from several SS genes in chimaeric constructs has shown that these regions can confer light-inducible expression upon genes not normally associated with this type of gene regulation (Broglie et al., 1984; Herrera-Estrella et al., 1984; Facciotti et al., 1985). The use of chimaeric genes containing deletion-derivatives of these 5'-untranslated regions has led to the definition of regions involved in SS gene regulation. Not only has it been shown that -400 bp of 5'-untranslated sequence from pea SS genes are sufficient to confer light-inducibility and organ-specific expression on a chimaeric gene (Nagy et al., 1985), but a 33 bp sequence

surrounding the TATA box has been identified as essential for light-inducibility (Morelli *et al.*, 1985). The enhancer-like element described by Timko *et al.* (1985b) and Fluhr and Chua (1986) is involved in phytochrome-induced transcription, white-light regulation and organ-specific expression of the SS genes.

The way in which these 5'-untranslated regions are affected by trans-acting factors is not known. The mode of action of phytochrome and the putative blue-light receptor in modulating transcription is also not known. The isolation of different members of SS multi-gene families from several plant species and a comparison of their 5'-untranslated regions may reveal sequences related to the relative degree of expression of those genes. The identification of trans-acting factors that affect SS gene regulation, and the development of an *in vitro* transcription system in which these factors can be assayed, should lead to a greater understanding of the control mechanisms involved in photomorphogenesis.

3. THE MAJOR CHLOROPHYLL a/b BINDING PROTEIN OF PHOTOSYSTEM II

3.A Function

The ATP and NADPH required during the carbon-fixing stages of photosynthesis are generated in the chloroplast by the passage of electrons through an electron transport chain located within the thylakoids, coupled to the formation of a proton gradient across the thylakoid bilayer. Incident light energy is collected primarily by the light-harvesting complexes LHC I and LHC II, and transmitted as excitation energy to the reaction centres of photosystem I (PS I) and photosystem II (PS II) respectively (for review see Bennett, 1983). Both the light-harvesting complexes and the photosystem reaction centres are located within the thylakoid membrane. The excitation energy derived from harvested light is used to oxidise water molecules from which oxygen is evolved and electrons generated. The electrons, so derived, are passed down the electron transport chain, from the PS II reaction centre to PS I. Once an electron reaches PS I, a further burst of excitation energy, provided chiefly from the PS I light-harvesting pigment bed, is fed into the chain. The production of both ATP and NADPH therefore involves an input of energy from both LHC I and LHC II. The major component of LHC II is the major chlorophyll a/b binding protein (LHCP). This protein is the most abundant thylakoid membrane protein (Bennett, 1983).

The LHCP is synthesised as a soluble precursor, consisting of an amino-terminal transit peptide linked to

the mature polypeptide, on free cytoplasmic ribosomes (Apel and Kloppstech, 1978; Schmidt et al., 1981). This precursor is post-translationally transported to the chloroplast where the transit peptide is proteolytically cleaved to yield the mature LHCP (Schmidt et al., 1981). The exact cleavage point of the transit peptide is not known but is either at Met^{34} or Arg^{35} in petunia (Dunsmuir, 1985), and at either Met^{37} or Arg^{38} in pea (Timko et al., 1985). The mature LHCP's associate with chlorophyll a and b and become integrated into the thylakoid membrane; the sequence of these events is not known. Once in the chloroplast, it is assembled into LHC II along with three or four other polypeptides. The major LHCP binds both chlorophyll a and b in LHC II and orientates them such that they are able to transmit their excitation energy to PS II (Bennett, 1983).

A short region of the LHCP protrudes from the membrane. This region of the protein is involved in the stacking of thylakoid membranes into grana. This same region of the protein can be reversibly phosphorylated by protein kinase and protein phosphatase enzymes located within the thylakoid membrane (Allen, 1983).

If the level of incident radiation is low, such as where leaves are shaded by other leaves, the level of PS I excitation is relatively greater than the level of PS II excitation. PS I is excited by longer wavelenght light than PS II; in the shade of leaves, photons of longer wavelength predominate. This situation is known as State I. In bright light, however, the level of excitation of PS II is greater than that of PS I; this is known as State II. In State II the LHCP is phosphorylated, resulting in its physical

dissociation from PS II. This results in an increase in the excitation of PS I with respect to PS II and leads to a more even distribution of the excitation energy, derived from harvested light, between the two photo-systems. This whole process is reversible (Bennett, 1983).

The LHCP therefore plays an important role in harvesting light, driving electron transport, stacking of thylakoid membranes into grana, and is involved in State I-State II transitions as the composition of the incident radiation changes.

3.B Gene structure

In this Section, the genes encoding the major chlorophyll a/b binding protein (termed AB genes) will be discussed. The structure of the AB genes and their organisation in the plant genome will be reviewed. The possible functions of the highly conserved regions of the polypeptides from various plant species will also be considered.

The genes encoding the major chlorophyll a/b binding protein are present within the nucleus of the plant cell (Kung et al., 1972) and are organised into multi-gene families in all plant species so far examined, namely pea (Coruzzi et al., 1983); Arabidopsis (Leutwiler et al., 1986); petunia, barley and tobacco (Dunsmuir et al., 1983b) Lemna (Tobin et al., 1984) and wheat (Lamppa et al., 1985a). The number of family members varies between species, but ranges from four in Arabidopsis thaliana (Leutwiler et al., 1986) to sixteen in Petunia hybrida (Dunsmuir et al., 1983b). As with the SS multi-gene family, the AB genes are

clustered within the genome. Three of the four AB genes from Arabidopsis are located within an 11 kb restriction fragment (Leutwiler et al., 1986). The multi-gene family is inherited as a single Mendelian unit in pea; the AB genes are located on chromosome 2 in this species (Polans et al., 1985).

Several AB genes have been isolated and characterised from different plant species. The cDNA sequences isolated are as follows: pea, two (Broglie et al., 1981; Coruzzi et al., 1983; S.M.Smith, unpublished results) and petunia, five (Dunsmuir et al., 1983b). The characterised genomic sequences encoding the major chlorophyll a/b binding protein are as follows: Arabidopsis, three, (Leutwiler et al., 1986); Lemna, one (Karlin-Neumann et al., 1985); pea, three (Cashmore, 1984; Timko and Cashmore, 1983; Timko et al., 1985; this work) petunia, ten (Dunsmuir et al., 1983b; Dunsmuir, 1985) and wheat, one (Lamppa et al., 1985a). Only one of the eighteen isolated AB genomic sequences contains an intron; namely the Lemna AB gene. The presence of this intron is not only uncharacteristic for AB genes, but resembles an insertion element in its structure (Karlin-Neumann et al., 1985). The presence of this intron in only two of the twelve Lemna AB genes (Karlin-Neumann et al., 1985), along with the presence of an extra intron in some SS genes (Dean et al., 1985b; Mazur and Chui, 1985) is an indication that introns may be lost or gained during the evolution of a multi-gene family.

The members of the AB multi-gene family from petunia can be divided into five sub-families using the sequence divergence within their coding and 3'-untranslated regions (Dunsmuir et al., 1983b). Different members of the

multi-gene family encode slightly different polypeptides (Dunsmuir *et al.*, 1983b; Dunsmuir, 1985). Differences in the aminoacyl sequence of the chlorophyll a/b binding protein occur between species as well as within species. Despite these sequence differences there are two regions of the polypeptide that are perfectly conserved within the petunia multi-gene family; these same regions are also highly conserved in the protein from different species, namely: *Arabidopsis* (Leutwiler *et al.*, 1986); *Lemna* (Karlin-Neumann *et al.*, 1985); pea (Cashmore, 1984; Timko *et al.*, 1985) and wheat (Lamppa *et al.*, 1985a). The first of these highly conserved regions consists of 28 amino acids in petunia and is located close to the amino terminus of the protein. The second homology block in petunia consists of 26 amino acids near the centre of the protein. Karlin-Neumann *et al.* (1985) have put forward a theoretical structure for the AB polypeptide from *Lemna* based on its amino acid sequence. This structure predicts a protein with the amino-terminus in the stroma followed by a trans-membrane region, and a short thylakoid lumen loop followed by a second trans-membrane domain. A second stromal domain follows this trans-membrane region, while the third trans-membrane domain is followed by the carboxy-terminus which faces into the thylakoid lumen. If the amino acid sequences encoded by all the characterised AB genes are compared to the sequence from *Lemna* and its predicted structure, several domains appear to be strongly conserved. The 28 amino acid amino-terminal region discussed above is located on the stromal side of the thylakoid membrane. The three predicted trans-membrane domains are also strongly conserved. The second trans-membrane region

and the stromal loop are composed of the amino acids that make up the second major homology block. The third trans-membrane region is also strongly conserved. The regions predicted to lie on the thylakoid lumen side of the membrane do not appear to be so highly conserved. The strongly conserved amino-terminal domain is most probably the region that undergoes reversible phosphorylation and is involved in granal stacking. The site of chlorophyll attatchment is not known.

3.C Gene regulation

The regulation mechanisms of AB genes show similarities to those encoding the SS of Rubisco. The amount of translatable mRNA for the LHCP is increased in light-grown as opposed to dark-grown plants (Apel and Kloppstech, 1978). Also, the response of AB genes to light varies among different species (Tobin and Silverthorne, 1985). In barley (Apel and Kloppstech, 1978) and lemna (Tobin, 1978; Tobin, 1981) the amounts of translatable mRNA in light-grown plants are much higher than in dark-grown ones. The amount of this mRNA declines when the plants are placed in the dark. In pea there are considerable amounts of AB transcript in etiolated seedlings (Gallagher and Ellis, 1982; Jenkins et al., 1983); these amounts are, however, increased upon illumination of the plant with white light. The observed changes in the amount of translatable AB transcripts are due to changes in physical RNA amount and not to changes in its translatability (Gallagher and Ellis, 1982; Gollmer and Apel, 1983; Stiekema et al., 1983b; Bennett et al., 1984). The level of AB gene regulation is

primarily transcriptional (Gallagher and Ellis, 1982; Gallagher et al., 1985; Silverthorne and Tobin, 1984). It has been observed in pea (Thompson et al., 1983), that the increase in AB transcript amount upon light-induction is paralleled by an increase in chloroplast DNA amount. Simpson et al. (1986b) have observed that AB genes are expressed strongly only in leaf mesophyll cells; this expression correlates strongly with the presence of chloroplasts. Further evidence of a link between AB gene expression and the presence of chloroplasts comes from the observation that AB gene transcription is blocked in the absence of functional chloroplasts (Batschauer et al., 1986); in contrast, the level of SS transcription is relatively unaffected. This work (Batschauer et al., 1986) indicates that nuclear transcription from AB genes is dependent upon a plastid-derived factor.

The transcription of AB genes is phytochrome-regulated (Apel, 1979; Tobin, 1981; Bennett et al., 1984; Nagy et al., 1986). AB gene transcription is induced by intensities of red-light much lower than those needed for the induction of SS genes (Kaufman et al., 1984). In contrast to SS gene expression, red-light alone is sufficient to induce the transcription of AB genes (Kaufman et al., 1985). There appears to be no blue-light regulation of AB transcription (Kaufman et al., 1985). The amount of AB transcripts is also modulated by cytokinin (Teyssendier de la Serve, 1985; Funckes-Shippy and Levine, 1985).

The abundance of AB transcripts in etiolated pea seedlings, but the absence of the AB polypeptide, is thought to reflect turnover of the protein in the absence of

chlorophyll (Bennett, 1981). Since chlorophyll cannot accumulate in the absence of the AB protein, it is desirable for the protein to be present at the onset of chlorophyll synthesis.

The relative degrees of transcription of different members of AB multi-gene families have been shown to differ in several plant species. In Arabidopsis, transcription from one of the three AB genes is barely detectable in light-grown plants whilst the other two genes are transcriptionally active (Leutwiler et al., 1986). Dunsmuir et al. (1983b) have demostrated differential expression of the petunia AB multi-gene family. In wheat, it is proposed that not only are different AB genes expressed to different degrees, but different AB genes are expressed at different stages of leaf development (Lamppa et al., 1985a).

Recent work involving chimaeric constructs of the 5'-untranslated regions from AB genes and reporter genes has shown that 400 bp of sequence directly 5' to the coding region of an AB gene from pea are sufficient to direct light-inducible and cell type-specific expression of the chimaeric gene (Simpson et al., 1985; Simpson et al., 1986a). The expression of a wheat AB genomic sequence in transgenic petunia and tobacco plants has been recently reported (Lamppa et al., 1985b) This gene contained 4.4 kb of sequence 5' to the transcription initiation site. The regions involved in the observed light-regulation and organ-specific expression of this gene have not been further defined.

It has been noted (Timko et al., 1985) that two AB genes isolated from pea contain sequences homologous to the

viral enhancer core region. Similar sequences are present in pea SS genes and are thought to be involved in conferring high levels of expression on these genes (Fluhr et al., 1986).

The isolation and characterisation of further AB genes, and an analysis of the degrees of expression of different gene-family members, may lead to a greater understanding of the sequences involved in the regulation of plant gene expression. The characterisation of several AB genes may also enable questions to be answered concerning the role of AB gene sub-families within the AB multi-gene family.

4. PLANT TRANSFORMATION

4.A Introduction

This review of plant transformation will be divided into five Sections which will be previewed in this Section (Section I.4.A). In Section I.4.B the agrobacteria will be discussed in relation to their role as tools for plant gene manipulation. The organisation of the Ti-plasmids of A.tumefaciens will be considered, as will the molecular basis and host-range of Agrobacterium-induced tumours. The mode of transfer of DNA from bacterium to plant will also be considered. Section I.4.C will follow up the work reviewed in Section I.4.B by discussing the modifications to Ti-plasmids that have enabled their use as vectors for the introduction of foreign genes into plants. The various vector systems presently available will be reviewed and the modes of selection for transformed plant cells will be discussed. In Section I.4.D methods of plant transformation that do not involve agrobacteria, namely viral vectors, direct gene transfer and pollen-assisted gene transfer will be discussed. The final Section (Section I.4.E) will review the literature on the foreign genes that have been expressed in plants. The advances made towards an understanding of gene regulation by the use of plant transformation, the organ and tissue-specific expression of introduced genes and the organellar targeting of proteins derived from introduced genes will be discussed. The expression of viral genes and herbicide-resistance genes in transformed plants will also be reviewed. The present limitations of plant transformation technology will be considered.

4.B The Ti-plasmids of A.tumefaciens

Gram-negative soil bacteria of the genus Agrobacterium are the causative agents of the plant diseases crown gall (Smith and Townsend, 1907) and hairy root (Moor et al., 1979). Neoplastic crown gall tumours are caused by infection of wounded plant tissue by A.tumefaciens; hairy root disease is caused by infection with A.rhizogenes. A.rhizogenes and their Ri-plasmids will not be covered in detail here (for review see Bevan and Chilton, 1982). At least 643 plant species from 331 genera are susceptible to crown gall disease (Nester et al., 1984); these are mostly dicotyledonous plants, relatively few monocotyledonous plants are affected. The discovery that Agrobacterium-induced plant tumours continued to grow in vitro in the absence of bacteria (White and Braun, 1942), gave rise to the theory that a tumour-inducing principle was transferred from the bacterium to the plant cell.

The presence of unusual amino acid derivatives, termed opines, in tumours incited by A.tumefaciens, and the ability of agrobacteria to degrade these opines, was further evidence of an important link between A.tumefaciens and crown gall (Petit et al., 1970). The first opines characterised were octopine and nopaline (Petit et al., 1970); several other opines have since been studied (Firmin and Fenwick, 1978; Ellis and Murphy, 1981; Petit et al., 1983; Chilton et al., 1984). Petit et al. (1970) showed that the type of opine produced by the tumour is dependent upon the strain of A.tumefaciens that induced the tumour. These

authors also showed that agrobacteria could use these opines as a sole source of carbon, nitrogen and energy. The type of opine that a particular strain of Agrobacterium can degrade is dependant upon the type of tumour it induces. Strains that induce tumours that synthesise octopine can degrade octopine but not nopaline and vice-versa.

The observation that avirulent strains of A.tumefaciens could become virulent upon co-inoculation of a wounded plant with virulent and avirulent bacteria (Kerr, 1969; Kerr, 1971) furthered the idea of an infectious entity as the cause of crown gall. The discovery of giant (~200 kb), low-copy number (1-2 per cell), plasmids in pathogenic strains of A.tumefaciens (Zaenen et al., 1974) led to the conclusion that virulence was associated with these plasmids (Van Larebeke et al., 1974; Van Larebeke et al., 1975; Watson et al., 1975). The virulence plasmids of A.tumefaciens are referred to as Ti- (tumour-inducing) plasmids whilst those of A.rhizogenes are referred to as Ri- (root-inducing) plasmids. The ability of agrobacteria to degrade specific opines is determined by their Ti-plasmids (Petit et al., 1970; Van Larebeke et al., 1974). It was later shown that the type of opine that a tumour synthesised was also Ti-plasmid dependent (Bomhoff et al., 1976; Montoya et al., 1977). As a consequence of these observations, the Ti-plasmids were concluded to be responsible for tumour induction, opine synthesis and opine catabolism. This conclusion led to the speculation that part, or all, of the Ti-plasmid was transferred from the bacterium to the plant cell. The Ti-plasmid was therefore considered as a suitable candidate for the tumour-inducing principle proposed by

authors also showed that agrobacteria could use these opines as a sole source of carbon, nitrogen and energy. The type of opine that a particular strain of Agrobacterium can degrade is dependant upon the type of tumour it induces. Strains that induce tumours that synthesise octopine can degrade octopine but not nopaline and vice-versa.

The observation that avirulent strains of A.tumefaciens could become virulent upon co-inoculation of a wounded plant with virulent and avirulent bacteria (Kerr, 1969; Kerr, 1971) furthered the idea of an infectious entity as the cause of crown gall. The discovery of giant (~200 kb), low-copy number (1-2 per cell), plasmids in pathogenic strains of A.tumefaciens (Zaenen et al., 1974) led to the conclusion that virulence was associated with these plasmids (Van Larebeke et al., 1974; Van Larebeke et al., 1975; Watson et al., 1975). The virulence plasmids of A.tumefaciens are referred to as Ti- (tumour-inducing) plasmids whilst those of A.rhizogenes are referred to as Ri- (root-inducing) plasmids. The ability of agrobacteria to degrade specific opines is determined by their Ti-plasmids (Petit et al., 1970; Van Larebeke et al., 1974). It was later shown that the type of opine that a tumour synthesised was also Ti-plasmid dependent (Bomhoff et al., 1976; Montoya et al., 1977). As a consequence of these observations, the Ti-plasmids were concluded to be responsible for tumour induction, opine synthesis and opine catabolism. This conclusion led to the speculation that part, or all, of the Ti-plasmid was transferred from the bacterium to the plant cell. The Ti-plasmid was therefore considered as a suitable candidate for the tumour-inducing principle proposed by

White and Braun (1941). The presence of Ti-plasmid DNA in plant cells was initially demonstrated by renaturation studies on DNA isolated from plant tumours (Chilton et al., 1977); this work indicated that a defined region of the Ti-plasmid was present in the transformed plant cells. Southern blot analysis of plant tumour-DNA has been used to confirm this observation (Thomashow et al., 1980a; Thomashow et al., 1980b; Willmitzer et al., 1980; De Beuckeleer et al., 1981). The specific part of the Ti-plasmid transferred to the plant cell is known as the T-region (transferred region) when associated with the Ti-plasmid and the T-DNA when transferred to the plant cell (for reviews see Bevan and Chilton, 1982; Van Montagu and Schell, 1982; Nester et al., 1984). T-DNA transfer to the plant cell occurs within only a few hours of the cultivation of plant cells with agrobacteria (Horsch et al., 1985; Virts and Gelvin, 1985). The T-DNA integrates into the nucleus of the plant cell and not into either the chloroplast or mitochondrial DNA (Chilton et al., 1980; Willmitzer et al., 1980). Within the nuclear DNA the T-DNA is organised in nucleosomes (Schafer et al., 1984). Recent work has suggested that the T-DNA may integrate into the chloroplast genome at a very low frequency (De Block et al., 1985). The site of integration of the T-DNA within the nuclear DNA has been investigated (Zambryski et al., 1980; Ursic et al., 1983) and is believed to occur randomly in either single copy or repeat-DNA (Andre et al., 1986; Chyi et al., 1986; Teeri et al., 1986). However, the majority of T-DNA insertion events characterised show insertion into repeat-DNA.

Initial characterisation of Ti-plasmids was

performed by restriction enzyme analysis to obtain physical maps (Chilton *et al.*, 1978a; Depicker *et al.*, 1980; De Vos *et al.*, 1981). From the restriction maps of octopine and nopaline Ti-plasmids, it was evident that these two Ti-plasmid types differ in several ways, yet share regions of homology (Chilton *et al.*, 1978b; Depicker *et al.*, 1978; Drummond *et al.*, 1978; Engler *et al.*, 1981), such as the core region of the T-DNA which encodes the tumour-induction genes (Depicker *et al.*, 1978; Drummond *et al.*, 1978; Joos *et al.*, 1983; Willmitzer *et al.*, 1983). Transposon mutagenesis has been employed by many researchers to produce functional maps of the octopine and nopaline Ti-plasmids (Holsters *et al.*, 1980; Ooms *et al.*, 1980; Garfinkel and Nester, 1980; Ooms *et al.*, 1981; Garfinkel *et al.*, 1981). From these functional maps, various regions of the octopine and nopaline Ti-plasmids have been defined; these include, other than the T-region, the virulence (*vir*) -region, the opine catabolism region, the origin of replication (*ori*) and the conjugal transfer region (*tra*). The *vir*-region is involved in the transfer of the T-region from the Ti-plasmid to the plant cell (Nester *et al.*, 1984); it is also implicated in specifying *Agrobacterium* host-range (Knauf *et al.*, 1984). The organisation and mode of action of the *vir*-region on the T-region is discussed later in this Section in relation to the mechanism of DNA transfer to the plant cell. The loci involved in opine catabolism and conjugal transfer are induced only in the presence of opines (see Lichtenstein and Draper 1985). The induction of conjugal transfer genes by opines accounts for the observation by Kerr (1969) that the tumour-inducing principle, now know to be the Ti-plasmid, is

transferred from virulent to avirulent strains of Agrobacterium in the presence of tumour tissue, the site of opine synthesis.

The T-regions of octopine and nopaline plasmids differ in both their physical arrangement and in some of their encoded functions. The common features of the octopine and nopaline T-regions will be considered first followed by a consideration of their differences. The core region of T-DNA contains 6 loci encoding transcripts designated 1, 2, 4, 5, 6a and 6b (Willmitzer et al., 1983). Some of these loci are essential for oncogenicity (Depicker et al., 1978); all are expressed in plant cells (Willmitzer et al., 1983). Three loci involved in tumour induction and maintenance have been identified as well as one locus involved in tumour morphology (Garfinkel et al., 1981; Leemans et al., 1982; Joos et al., 1983; Willmitzer et al., 1983). Mutation of the tumour morphology (tml) locus, possibly encoding transcript 6b, leads to abnormally large tumours (Garfinkel et al., 1981). The tmr locus, encoding transcript 4 (Garfinkel et al., 1981), is involved in cytokinin biosynthesis (Akiyoshi et al., 1984; Buchmann et al., 1985). Mutations in the tmr locus lead to "rooty" tumours (Garfinkel et al., 1981; Joos et al., 1983; Willmitzer et al., 1983). Transcripts 1 and 2 of the T-region are encoded by the loci tms1 and tms2 respectively. A mutation in either tms1 or tms2 leads to "shooty" tumours (Garfinkel et al., 1981; Joos et al., 1983; Willmitzer et al., 1983). These loci are involved in auxin biosynthesis (Inze et al., 1984; Thomashow et al., 1984; Black et al., 1986; Thomashow et al., 1986). The tms1 and tms2 loci encode enzymes involved in consecutive steps in

the synthesis of indole-3-acetic acid (van Slogteren et al., 1984a; Thomashow et al., 1986). Tumour induction is therefore due to phytohormone imbalance within the plant cell caused by transcripts 1, 2 and 4; transcripts 5, 6a and 6b are not essential for tumour formation (Garfinkel et al., 1981; Joos et al., 1983).

Messens et al. (1985) have identified transcript 6a as being involved in octopine and nopaline secretion. No confirmed function has been ascribed to either transcript 5 or 6b. However, Joos et al. (1983) indicate that transcript 5 may be involved in host-range determination and Garfinkel et al. (1981) indicate that transcript 6b may be involved in determining the size of tumours. The observation by Joos et al. (1983), that T-DNA mutations leading to the loss of all three tumour induction genes (1, 2 and 4) dose not affect T-DNA transfer, led to the development of disarmed Ti-plasmid vectors for plant transformation.

The transcripts from T-DNA genes expressed in plant cells are polyadenylated (Willmitzer et al., 1982; Dhaese et al., 1983) and their expression in plant cells is α-amanitin-sensitive. Transcription of T-DNA genes in plants is therefore RNA polII-dependent (Willmitzer et al., 1981). It is therefore surprising that four of the genes in the core region of the T-region are also expressed in agrobacteria (Schroder et al., 1983). The expression of these genes in the Agrobacterium probably assist in the initiation of tumour induction. In addition to the phytohormones produced by the T-DNA genes, agrobacteria produce the auxin, IAA (Lui et al., 1982) and the cytokinin, trans-zeatin (Beaty et al., 1986). The genes encoding these

phytohormones are not transferred to the plant cell and are also thought to be involved with the initiation of tumour formation.

The greatest difference between the T-regions of octopine and nopaline Ti-plasmids is the division of the octopine plasmid T-region into two separately transferred segments, as compared to the single T-region found in nopaline plasmids. The two T-regions of octopine type plasmids are known as T_L and T_R; the nomenclature is consistent with their relative positions on the Ti-plasmid (Thomashow et al., 1980a; De Beuckeleer et al., 1981). The T_L-region also contains the octopine synthase (ocs) gene. The ocs gene encodes transcript 3 (Schroder et al., 1981). This gene has been sequenced (De Greve et al., 1983) and contains all the sequences required for expression in plant cells (Koncz et al., 1983). Analyses of T_L-DNA transcripts in plant cells show that T_L-DNA encodes eight polyadenylated transcripts: 1, 2, 4, 5, 6a, 6b and 7 (Gelvin et al., 1982; Willmitzer et al., 1982; Gielen et al., 1984). Transcripts 1, 2, 3, 4, 5, 6a and 6b are within the core region and have already been discussed, the function of transcript 7 is still not known. Mutations in regions of the T_L-DNA outside the core region do not affect tumour formation or morphology (Joos et al., 1983). The T-DNA region from nopaline Ti-plasmids contains, as well as the core region loci, the nopaline synthase (nos) gene (Holsters et al., 1980; Willmitzer et al., 1983). The nos gene has been sequenced by Depicker et al. (1982) and its activity analysed by An et al. (1986). The T_R-DNA of octopine Ti-plasmids encodes five polyadenylated transcripts in plant cells and does not

incite tumours in the absence of T_L-DNA. Two of the transcripts are believed to be involved in mannopine and agropine synthesis (Karcher *et al.*, 1984); the function of the other three transcripts is unknown. T_L-DNA and T_R-DNA are transferred to the plant cell independently. Transformed cells may therefore contain only T_L-DNA or T_L- and T_R-DNA; these cells will be tumourous. Alternatively, the transformed cells may contain only T_R-DNA (Czako and Marton, 1986) in which case they will not be tumourous.

In summary, the T-DNA of nopaline plasmids is known to encode 13 polyadenylated trascripts (Willmitzer *et al.*, 1983); the T_L-DNA from octopine plasmids encodes 8 polyadenylated transcripts (Leemans *et al.*, 1982), and 6 of these are conserved between both types of plasmids (Garfinkel *et al.*, 1981). The T_R-DNA from octopine plasmids encodes 5 polyadenylated transcripts, none of which are present in nopaline plasmids (Karcher *et al.*, 1984). No T-DNA-encoded genes containing introns have been observed (Gielen *et al.*, 1984).

The degree of expression of the various T-DNA genes have been studied and the amounts of transcripts shown to vary for different genes (Willmitzer *et al.*, 1983; Karcher *et al.*, 1984). Gene 5 from octopine plasmid T_L-DNA has been shown to be expressed only in plant tissues containing high auxin levels (Koncz and Schell, 1986). The high constitutive levels of *ocs* and *nos* gene expression in transformed tissues (Otten *et al.*, 1981; DeGreve *et al.*, 1983) has led to the use of their promoters for high level expression of chimaeric genes in transformed plant cells; this will be discussed in more detail in Section I.4.D.

Ti-plasmid-induced tumours occasionally produce shoots. These spontaneous shoots can be recognised as transformed since they contain opines. The regenerative capacity of some transformed plant cells was shown to be associated with the loss of some T-DNA sequences (Yang and Simpson, 1981; DeGreve et al., 1982). These shoots produced roots and developed into normal plants. Wullems et al. (1981b) observed that spontaneous shoot production often led to shoots that displayed the characteristics of high cytokinin levels, namely, they were unable to form roots, formed lateral shoots and showed no apical dominance. This unusual growth pattern was due to the loss of the auxin-producing loci tms1 and tms2. These shoots were observed to grow into mature, fertile plants if they were grafted onto non-transformed root systems. The insertional inactivation of the tmr locus of the T-DNA (Barton et al., 1983) led to "rooty" tumours from which plants could occasionally be regenerated. Octopine and nopaline plasmids differ in the frequency with which they spontaneously produce shoots. The tms loci of nopaline Ti-plasmids are weaker than those of octopine plasmids. Loss of the tmr locus from the T-DNA of nopaline plasmids therefore leads to transformed plant cells whose hormone balance is not altered drastically by the weak tms loci. Cells containing tms-deleted nopaline T-DNA therefore regenerate more often than cells containing tms-deleted octopine T-DNA (Nester et al., 1984).

T-DNA can be methylated in plant cells (Gelvin et al., 1983). This methylation may lead to the inactivation of the tumour-inducing loci and the subsequent formation of

shoots from the tumour. These shoots can be regenerated into normal plants (Amasino et al., 1984). Demethylayion of the T-DNA by treatment with 5-azacytidine can re-initiate tumourous growth. The grafting of spontaneous shoots, that show no octopine synthase activity, onto normal stems can lead to the reactivation of the octopine synthase gene (van Slogteren et al., 1984b).

From these observations, it became apparent that T-DNA transfer to the plant cell did not rely on oncogenicity. However, if Ti-plasmids were to be used as plant transformation vectors, a means of regenerating whole plants was required that did not rely on the rare spontaneous deletion of part of the T-DNA. Ti-plasmids with disarmed T-regions provided the answer. The inactivation of transferred genes by methylation has not yet been overcome, and the degree of expression of introduced genes may vary considerably in individual transformed plants.

The numbers of T-DNA copies transferred to the plant genome can range from one to twenty (Merlo et al., 1980; Thomashow et al., 1980a). It has been noted that T-DNA sequences from both octopine and nopaline Ti-plasmids can be inserted into the same plant cell (Wullems et al., 1981a). More recent work has investigated the phenomena behind this observation. The work of Depicker et al. (1985) and Petit et al. (1986) indicates that T-DNA from different agrobacteria can be introduced into the same cell, upon co-infection by a mixed population of agrobacteria, to give multiple transformation of that cell. It was however noted (Depicker et al., 1985), that multiple copies of T-DNA within a cell are most likely due to the multiple transfer of T-DNA from

one bacterium rather than the multiple infection of one cell. The work of Depicker et al. (1985), Petit et al. (1986) and de Frammond et al. (1986) indicates that transfer of two independent T-DNA sequences from within the same bacterium occurs at a higher frequency than does transfer of the same two sequences from different bacteria. These observations may be of relevance to the use of Ti-plasmids as gene vectors when high copy-number insertions of foreign sequences, or high levels of expression are required. The observation that DNA inserted into the T-region of Ti-plasmids (Garfinkel et al., 1981; Otten et al., 1981) was transferred to the plant chromosome along with the rest of the T-DNA was a further indication that Ti-plasmids could be used as gene transfer vectors for plants.

The extreme ends of the T-region of Ti-plasmids are defined by 25 bp direct imperfect repeats (Yadav et al., 1982). These repeats and their flanking regions, known collectively as T-DNA borders, are responsible for the transfer of the T-DNA, and any DNA located within it, to the plant cell (Simpson et al., 1982; Yadav et al., 1982; Holsters et al., 1983; Wang et al., 1984). Up to 50 kb of foreign DNA can be inserted into the T-region and transferred from the Ti-plasmid to the plant cell (Herrera-Estrella et al., 1983a). The right-hand copy of the 25 bp repeat is essential for T-DNA transfer which it can promote in the absence of the left-hand repeat; the left-hand 25 bp repeat is unable to promote T-DNA transfer from most Ti-plasmids in the absence of the right-hand repeat (Joos et al., 1983; Wang et al., 1984; Shaw et al., 1984). The left-hand repeat however shows some T-DNA

transfer ability from mini-Ti plasmids (Jen et al., 1986a; Jen and Chilton, 1986b). Mini-Ti plasmids are artificial Ti-plasmids containing only the T-region; the vir region is located on a separate replicon and functions in trans to the T-region (Jen et al., 1986a) This binary transfer system is discussed in greater detail below. The right-hand 25 bp repeat is directional in its action, promoting T-DNA transfer in a polar fashion (Wang et al., 1984). The 25 bp right-hand repeat promotes only weak T-DNA transfer when the flanking regions from the right hand side of the repeat are deleted (Peralta and Ream, 1985). This weak T-DNA transfer is enhanced by the presence of a 15 bp conserved sequence, termed overdrive, located to the right of the right-hand repeat (Peralta et al., 1986). The T-DNA 25 bp repeats are discussed in greater detail below in relation to the mechanism of T-DNA transfer.

The function of T-DNA transfer in nature is discussed by Tempe and Petit (1985), who note that the ability to degrade a particular opine, encoded by a Ti- or Ri-plasmid, is always associated with the ability to direct its synthesis in plant cells. Opine metabolism therefore provides an ecological niche and a selective advantage for agrobacteria possesing Ti- or Ri-plasmids.

The agrobacteria are divide into two classes according to their host-range. The broad host-range type of A.tumefaciens can induce tumours on a wide range of dicotyledonous plants whereas the limited host-range types are usually limited to grapevines (Thomson, 1986). The host-range of a strain of A.tumefaciens is defined in part by its Ti-plasmid (Thomashow et al., 1980a). The T-region of

Ti-plasmids that confer a limited host-range is divided into two separate regions termed T_A and T_B (Buchholz *et al.*, 1984a; Knauf *et al.*, 1984). T_A encodes the auxin-independence loci (transcripts 1 and 2), whilst T_B encodes the cytokinin-independence locus (transcript 4) and the opine secretion locus (transcript 6a) (Buchholz *et al.*, 1984a). Buchholz *et al.* (1984b) have extended the host-range of a limited host range Ti-plasmid by transferring the cytokinin-independence gene, *tmr*, from a broad host range Ti-plasmid into the T_B-region of the limited host range plasmid. This observation indicates that the cytokinin independance gene, *tmr*, is involved in the determination of host-range; similar conclusions were drawn by Hoekema *et al.* (1984a). This observation correlates with the observations that broad host range Ti-plasmids lacking the *tmr* locus incite tumours on fewer plant species (Garfinkel *et al.*, 1980; Garfinkel *et al.*, 1981; Ooms *et al.*, 1981). It has been noted by Knauf *et al.* (1984) that two loci from the virulence region of the Ti-plasmid, *virA* and *virC*, may play a role in host range limitation. Introduction of these two loci into limited host range strains increases the host-range of these strains (Knauf *et al.*, 1984). The host-range limitation of some *Agrobacterium* strains may also involve chromosomal loci (Garfinkel and Nester, 1980) and plant cell factors involved in plant cell-agrobacterium binding (Nester *et al.*, 1984). Host-range determination is therefore a complex and at present poorly understood subject.

It is possible that the ability of an *A.tumefaciens* strain to induce a tumour on a particular plant species is

dependent upon the susceptibility of the plant species to perturbation of its phytohormone levels. The absence of tumour formation is therefore not an indication that DNA transfer has not occurred. This conclusion may be a favourable one from the point of view of modifying plants that were thought to be immune to A.tumefaciens-mediated transformation.

The vir region of Ti-plamids is responsible for the transfer of the T-DNA to plant cells (Garfinkel and Nester, 1980: Klee et al., 1983). None of the vir loci are transferred to the plant cell; they are expressed in the bacterium (Das et al., 1986). From transposon-mutagenesis studies on the vir region of the octopine Ti-plasmid, twelve loci involved in virulence have been identified by complementation analysis: these loci are named as follows: virA to virE (Klee et al., 1982; Klee et al., 1983). The virB locus has been shown to contain six units of gene expression, designated virB I to virB VI (Iyer et al., 1982). Two further loci, virG and pinF (Stachel and Nester, 1986) have been identified. Transposon-mutagenesis of nopaline Ti plasmids has as yet revealed only six complementation groups (Lundquist et al., 1984). The mutations induced by transposon-insertion into the vir loci most often result in complete loss of virulence. However some mutations lead to delayed tumour induction (Iyer et al., 1982) or altered host-range (Klee et al., 1983). The expression of several of the vir loci has been shown to be plant-inducible. (Okker et al., 1984: Stachel et al., 1986a). Several phenolic compounds of plant origin are able to cause this induction (Bolton et al., 1986). Two of these,

acetosyringone and α-hydroxyacetosyringone (Stachel et al., 1985), are produced only by wounded plant cells.

It has been show that the loci virB, virC, virD, virE and pinF are only expressed upon induction by plant cell extracts (Das et al., 1986; Stachel et al., 1986a); pinF (plant-inducible) is not involved in virulence but is associated with the vir region. The loci virA and virG are expressed constitutively in the agrobacterium cell (Das et al., 1986). The level of virG expression is however further induced by plant cell factors (Stachel et al., 1986a).The expression of the virB, virC, virD, virE loci is not only plant-inducible but requires the presence of the virA and virG gene products (Das et al., 1986). Mutation in any of the virA, virB, virD or virG loci lead to avirulence whilst mutation of either the virC or virE locus leads to attenuated virulence (Stachel and Nester, 1986). The virE locus can be complemented extracellularly (Otten et al., 1985), implying the production of a diffusible product by the virE locus. This observation and the work of Gardner and Knauf (1986) led these authors to the conclusion that Agrobacterium-mediated plant transformation occurs in two stages. The first stage involves the virA, virB, virC and virD loci but not virE: the second stage requires virE. The role of virE was speculated by Gardner and Knauf (1986) to be involved in the integration of the T-DNA into the plant genome. Stachel and Nester (1986) postulate that virB or virD may be involved in the synthesis of a T-DNA intermediate for transfer to the plant cell, virB, virE, virG and pinF may encode structural proteins, and virA, virC and virD may specify enzymatic functions.

Two chromosomal virulence loci, chvA and chvB have been identified (Douglas et al., 1985); these loci are constitutvely expressed and are not induced by plant cell factors. The way in which all the virulence loci act in concert to promote the transfer of the T-region from the Ti-plasmid to the plant chromosome is yet to be resolved.

The observations that the vir region can promote transfer of the T-region to plant cells when the two regions are located on separate plasmids (deFrammond et al., 1983; Hoekema et al., 1983), or when the T-region is located on the agrobacterium chromosome (Hoekema et al., 1984b), and the generality of the vir regions between various Ti-plasmids (Hoekema et al., 1984c; Hooykaas et al., 1984) has led to the development of several vector systems for plant transformation; these will be discussed in Section I.4.E. The identification of part of the vir-region from an octopine Ti-plasmid as being involved in enhancing virulence of a nopaline Ti-plasmid may be of significance to the development of efficient vectors for plant transformation (Otten et al., 1985).

The isolation of T-DNA circles from agrobacteria co-cultivated with plant cells (Koukolikova-Nicola et al., 1985) led to speculation that these molecules may be intermediates in the T-DNA transfer process. The circularisation is plant cell-induced (Machida et al., 1986; Stachel et al., 1986a) and requires the presence of the vir region. (Koukolikova-Nicola et al., 1985; Machida et al., 1986). The virB and virC loci, both inducible by plant cell extracts (Stachel et al., 1986a), have been identified as being involved in the formation of the T-DNA circles

(Alt-Moerbe et al., 1986). The junction of the T-DNA circles occurs in the 25 bp direct repeats only when the repeats are in the same orientation (Machida et al., 1986). A more recent observation by Stachel et al. (1986b) indicates that the T-DNA circles seen in acetosyringone-induced agrobacteria are not the transfer intermediate. Stachel et al. (1986b) have identified single-stranded copies of the T-DNA, T-strands, that they propose are transferred to the plant cell by bacterial-plant cell conjugation. The polarity of T-strands is dependent on the orientation of the 25 bp repeats and involves the introduction of single-strand nicks within the T-DNA border sequences. DNA synthesis is then primed from the free 3'-OH group generated by the single-stranded nick adjacent to the right hand 25 bp repeat. Synthesis of this new strand displaces the existing nicked strand to generate the polar T-strand. It is proposed that T-DNA circles are formed by nick-induced recombination between the T-DNA 25 bp repeats. These circles are therefore not believed to be involved in T-DNA transfer but are merely by-products of T-strand synthesis. The actual mechanisms of T-strand transfer to the plant cell, the steps involved in its movement to the nucleus and its integration within the plant chromosome are not yet fully understood.

An understanding of the structure and function of the T- and vir-regions has led to the development of Ti-plasmids as vectors for the routine transfer of foreign genes into plant cells. A greater understanding of host-range, and T-strand transfer may lead to the development of further improved vectors.

4.C Ti-plasmid-based vectors for plant transformation

The development of Ti-plasmid-based vectors for the introduction of foreign genes into plant cells arose after several key observations. Firstly, that the T-DNA was specifically integrated and stably maintained in the chromosomes of plant cells (Chilton et al., 1980; Willmitzer et al., 1980); secondly, that transposons or insertion elements inserted into the T-DNA were co-transferred with the T-DNA to the plant cells (Ooms et al., 1983; Garfinkel et al., 1981); thirdly, that deletion of the tumour induction genes did not affect T-DNA transfer (Joos et al., 1983), and fourthly that entire plants could occasionally be regenerated from tumour tissue (Barton et al., 1983; Wullems et al., 1981b). Further observations that have led to the refinement of Ti-plasmid vectors are the ability of the vir-region to act in trans upon the T-region (de Frammond et al., 1983; Hoekema et al., 1983), the ability to select transformed cells in the absence of hormone-independent growth (see below, this Section), and the ability to regenerate plants transformed with disarmed Ti-plasmids (De Block et al., 1984). The background to these observations and the molecular basis of T-DNA transfer have been discussed above (Section I.4.A). The currently available Ti-plasmid vectors fall into two categories; those that rely on homologous recombination to introduce the gene destined for the plant chromosome into the T-region of a Ti-plasmid within A.tumefaciens. and the binary vectors that rely on the trans-acting vir-region to transfer the T-region from a second replicon to the plant chromosome. The recombination vectors will be considered first followed by the binary vectors; the selection of plant cells transformed by these

vectors will also be considered.

Due to the large size of the Ti-plasmids (~200 kb), *in vitro* manipulations to introduce DNA into the T-region are technically difficult. The first approach towards overcoming this problem involved the subcloning of a T-DNA restriction fragment into an *E.coli* plasmid vector. The genes encoding yeast alcohol dehydrogenase and a bacterial kanamycin-resistance marker were ligated into the T-DNA fragment causing insertional inactivation of the *tmr* locus (Barton *et al.*, 1983). The resulting engineered T-DNA fragment was subcloned into a broad-host range vector and transformed into an *A.tumefaciens* strain harbouring a wild-type Ti-plasmid. Following a double homologous recombination event between the mature Ti-plasmid and the engineered T-DNA, the wild-type *tmr* locus was exchanged for the insertionally inactivated one. The broad-host range plasmid was then evicted by plasmid incompatibility and the modified Ti-plasmid selected on kanamycin. Induction of tumours on plants with this partially disarmed Ti-plasmid led to "rooty" type tumours from which some mature plants were regenerated. The ability to mobilize plasmids from *E.coli* to *A.tumefaciens* by conjugal transfer (Van Haute *et al.*, 1983) was exploited by Shaw *et al.* (1983) who demonstrated that an intermediate vector, capable of replication in both *E.coli* and *A.tumefaciens*, could be used to introduce genes into *Agrobacterium*, thus eliminating an extra cloning step into a broad-host range vector. This vector system also relied on homology-mediated recombination at two sites to transfer the gene destined for the plant chromosome from the intermediate vector into the Ti-plasmid.

Selection of these rare double recombinants was via antibiotic-resistance genes. These vector systems were refined still further by the introduction of pBR322 homologous sequences into the T-region of a Ti-plasmid (Koncz et al., 1984; De Block et al., 1984). The mobilisation of pBR322-derived vectors containing foreign DNA sequences into agrobacteria harbouring these modified Ti-plasmids led to homologous recombination between the pBR322-derived vector and the modified Ti-plasmid. Double or single recombination events resulted in the integration of the foreign DNA into the T-region. Single recombination events occurred at a higher frequency than did double recombination events and were found to be stable under antibiotic selection; this observation simplified the procedure of introducing DNA into the T-region in vivo. These vectors provided a simple method for the introduction of foreign genes into plants. In the vector system described by De Block et al. (1984), the Ti-plasmid was disarmed to enable the regeneration of plants from transformed cells. A further modification to this type of vector involved the replacement of the entire T-region, including T-DNA borders, of a Ti-plasmid with pBR322. The intermediate vector contains the T-DNA borders into which are inserted a selectable marker and the gene for transfer to the plant chromosome. Upon cointegration of the Ti-plasmid and the intermediate vector, a T-region, including border sequences, is reintroduced into the resident Ti-plasmid (Deblaere et al., 1985). The advantage of this system over the one devised by DeBlock et al. (1984) is the absence of pBR322 sequences between the T-DNA borders. In this system pBR322

sequences lie outside the T-DNA borders and are not transferred to the plant cell. In the system of De Block *et al*. (1984), the pBR322 sequence lies within the T-DNA borders resulting in the transfer of this DNA and the entire intermediate vector to the plant cell. The presence of these extra sequence complicates the analysis of the T-DNA within the plant genome.

The split-end vector system (SEV) of Horsch *et al*. (1984) relies on homologous recombination to introduce an intermediate vector containing the gene for transfer to the plant cell into a Ti-plasmid resident in *Agrobacterium*. The intermediate vector contains only the right T-DNA border, selectable markers, the *nos* gene and a region of homology to the T-region; no left T-DNA border is present. Upon introduction of this intermediate vector into a strain of *A.tumefaciens* harbouring an octopine Ti-plasmid, a single recombination event results in a cointegrate molecule that is selectable by its antibiotic resistance markers. The site of integration, determined by the region of homology to the resident Ti-plasmid, places the right T-DNA border of the intermediate vector between the left and right borders of the resident octopine Ti-plasmid. Upon cocultivation of *A.tumefaciens* and plant protoplasts, two possible T-strands can be transferred to the plant cells; a long strand, transferred from the right border of the octopine Ti-plasmid and containing the tumour induction genes, the *ocs* gene and the intermediate vector; or a short strand, transferred from the right border of the intermediate vector, omitting the tumour induction genes and the *ocs* gene. The only DNA

transferred by the short strand is the disarmed T-region and the sequences within it. The transformants derived from long strand transfer (~90%) are unable to regenerate into plants because of the tumour induction genes, whilst those derived from short strand transfer (~10%) can regenerate into intact plants. A refined version of SEV system has been described by Fraley *et al*. (1985) and employs a resident Ti-plasmid lacking the right T-DNA border sequence; only short strand transfer can therefore occur. This refined vector system improves the transfer efficiency 10-fold since all the transformed cells are able to regenerate into whole plants.

The alternative approach to reducing the size of the Ti-plasmid for *in vitro* manipulations is to divide the *vir* and T-regions between different replicons and exploit the *trans* activity of the *vir*-region on the T-region. These vectors are referred to as binary vectors. Several binary vector systems have been described (de Frammond *et al*., 1983; Hoekema *et al*., 1983; Bevan, 1984; An *et al*., 1985; Hoekema *et al*., 1985; Klee *et al*., 1985; Van den Eltzen *et al*., 1985a; Simpson *et al*., 1986a) in which the disarmed T-region is present on a broad-host range plasmid, capable of replication in both *E.coli* and *A.tumefaciens*. The gene for transfer to the plant cell is ligated into the T-region. The recombinants are selected in *E.coli*, by antibiotic resistance markers present on the vector, and transferred to *Agrobacterium* either by transformation (de Frammond *et al*., 1983) or by conjugation. All the systems described, other than those of Hoekema *et al*. (1985) and Klee *et al*. (1985), which are self transmissible, need to be mobilised into an *Agrobacterium* strain harbouring a T-DNA deleted Ti-plasmid

by a third replicon. Some binary vector systems are based around cosmid-T-DNA vectors (An et al., 1985; Klee et al., 1985; Van den Eltzen et al., 1985a); these vectors allow the cloning of large pieces (30-40 kb) of DNA into plant cells, or the recovery of this DNA from transformed plants. The binary-vector system of Bevan (1984) was used for the work contained in this thesis because of its availability and ease of use. This vector system is described in greater detail in Section III.5.A.

In order to obtain the full potential from the available plant transformation vectors it is essential to be able to differentiate between transformed and non-transformed cells. Since phytohormone-independance cannot be used as a selectable marker in cells transformed with disarmed Ti-plasmid vectors, drug-resistance markers have been developed that enable the identification of transformed cells. As bacterial drug-resistance genes do not function under the control of their own promoters in plant cells (Fraley et al., 1983), several chimaeric drug-resistance genes have been constructed that use the 5'- and 3'-regulatory regions of various constitutively-expressed plant genes to provide high levels of expression of the resistance marker in all plant cell types. The promoter sequences from the octopine and nopaline synthase genes have been shown to be constitutively expessed in transformed plants (Otten et al., 1981; De Greve et al., 1982). Selectable markers have been constructed that comprise of the regulatory regions of the nos gene linked to the neomycin phosphotransferase I (nptI) and the neomycin phosphotransferase II (nptII) genes from Tn601 and Tn5

respectively (Bevan et al., 1983; Fraley et al., 1983; Herrera-Estrella et al., 1983a). These markers confer resistance to aminoglycoside antibiotics, such as G418, kanamycin and neomycin, upon the cells harbouring the chimaeric genes. The level of resistance conferred by nos-nptI genes is however lower than that conferred by nos-nptII genes. The use of the nos promoter to express the bacterial chloramphenicol acetyltransferase gene (cat) and the methotrexate-resistant dihydrofolate reductase gene (Mtx^R DHFR) in plant cells, and to confer chloramphenicol and methotrexate-resistance on the transformed plant cells, has been described (Herrera-Estrella et al., 1983a; Herrera-Estrella et al., 1983b; De Block et al., 1984). De Block et al. (1985) have indicated that in some cases chloramphenicol-resistance is due to chloroplast transformation rather than nuclear transformation. Other chimaeric antibiotic-resistance genes have been constructed that confer resistance to the antibiotics hygromycin-B (van den Elzen et al., 1985b; Waldron et al., 1985) and bleomycin (Hille et al., 1986). Several other constitutive promoters have been used to express selectable markers, namely the 35S cauliflower mosaic virus (CaMV) promoter to express kanamycin and chloramphenicol-resistance genes (Pietrzak et al., 1986), the CaMV inclusion body gene promoter to express the nptII gene from Tn5 (Koziel et al., 1984), and the CaMV gene IV promoter to express this same gene (Paszkowski et al., 1984; Balazs et al., 1985; Paszkowski et al., 1986). The use of a divergent T-DNA promoter to express kanamycin and chloramphenicol-resistance genes in transformed plants has been demonstrated by Velten and Schell (1986). The

advantage of this divergent promoter is that other genes can be expressed in place of one of the marker genes; expression of antibiotic-resistance would indicate the expression of the second gene from the dual promoter.

The infection of plant cells with agrobacteria harbouring Ti-plasmid vectors is necessary to achieve T-DNA transfer. This process can involve either the co-cultivation of plant protoplasts with the agrobacteria (Marton et al., 1979; Fraley et al., 1984) or the infection of leaf discs (Horsch et al., 1985). An advantage of the protoplast method is the ability to select and regenerate individual transformed plant cells, whereas the leaf disc infection method provides a simpler and faster means of transforming plant cells.

The effective use of Ti-plasmid-derived vectors to transfer foreign genes into plant cells has been demonstrated in many laboratories; examples will be discussed in Section I.4.D. There are, however, several limitations to the system such as the number of monocotyledonous plants susceptible to Agrobacterium infection and the current inability to regenerate many plant species from tissue culture. To overcome some of these problems, alternative transformation techniques are being developed.

4.D Non-Agrobacterium-mediated plant transformation

Although Agrobacterium mediated transformation with Ti-plasmid vectors can be used as a reliable and relatively simple method for introducing foreign DNA into plant cells, it suffers from several dissadvantages. Firstly, the plant

to be modified must be susceptible to Agrobacterium infection. This limits the technique to most dicotyledonous plants and relatively few monocotyledonous plants, namely members of the Lilliaceae and Amaryllidaceae (Hernalsteens et al., 1984; Hooykaas-van Slogteren et al., 1984). Secondly, several cloning steps may be required to transfer the gene of interest from an E.coli cloning vector into the Ti-plasmid vector; and thirdly, the susceptible plant must be amenable to regeneration from tissue culture; again this precludes most monocotyledonous plants. To overcome these limitations, several other DNA transfer systems have been investigated. The first one considered here is direct transformation, where naked DNA is taken up by the plant cells and becomes integrated into the plant chromosomes; the second alternative to the Ti-plasmid-based vectors considered, is the use of virus-based vectors. The final method of plant cell transformation that will be considered is that of pollen-mediated DNA uptake.

The observation by Krens et al. (1982) that naked Ti-plasmid DNA could be transferred to tobacco protoplasts by polyethylene glycol/Ca^{2+}-mediated transformation, and that tumour induction ensued, was an early indication that direct transformation of plant cells was feasible. It was not known at the time whether the integration and expression of the tumour induction genes followed the same path as when transformation was mediated by Agrobacterium. It was however noted that DNA other than the T-DNA was integrated into the plant genome and that shoots, of the same morphology as those obtained from shooty tumours by Wullems et al. (1981b), could develop from some tumour clones. From

a consideration of more recent data, such as the prokaryotic organisation of the Ti-plasmid *vir* genes, and their lack of eukaryotic promoter sequences (Das *et al.*, 1986), it is evident that the presence of non T-DNA sequences in the plant cells and spontaneous shoot formation are due to integration of the Ti-plasmid sequences into the plant genome through non-homologous recombination, and not due to the influence of the *vir* region.

Non-homologous integration of foreign DNA into the genomes of tobacco protoplasts transformed by naked DNA of an *E.coli* plasmid vector has been reported (Paszkowski *et al.*, 1984; Balazs *et al.*, 1985). Protoplasts were transformed by the PEG/Ca^{2+} method of Krens *et al.* (1982), selected via a chimaeric kanamycin-resistance gene and regenerated into mature plants. The kanamycin-resitance marker was inherited in a Mendelian fashion by the R_1 generation. An alternative approach to this direct transformation technique has been described (Deshayes *et al.*, 1985), that uses liposome mediated transformation of tobacco protoplasts to introduce foreign DNA into their genomes. Liposome mediated transformation is reportedly 20 times as efficient as the PEG/Ca^{2+}-mediated method (Deshayes *et al.*, 1985). The foreign DNA appears to be integrated as concatermeric structures consisting of tandemly-repeated DNA fragments (Hain *et al.*, 1985; Potrykus *et al.*, 1985a). Similar observations have been made in transgenic mice, as described by Palmiter and Brinster (1985), and Xenopus (Etkin *et al.*, 1984). The site of integration of this exogenous DNA in mouse cells occurs within repetitive sequences (Kato *et al.*, 1986). The site of integration in

plant cells is as yet unknown. The techniques of direct transformation have been applied to protoplasts from monocotyledonous plants with the same success as observed in protoplasts from dicotyledonous plants. Lolium multiflorum (Italian rye grass) protoplasts have been stably transformed by direct transformation (Potrykus et al., 1985b), as have protoplasts of Triticum monococcum (Lorz et al., 1985). These two plant species are from the family Gramineae and are immune to Agrobacterium-mediated transformation. The commercial importance of the Gramineae as crop plants makes them attractive for modification by genetic engineering. The demonstration that exogenous DNA can be transferred by direct transformation to the chromosomes of species in the family Gramineae is an encouraging result. The problems of their tissue culture and regeneration are the next obstacles to be overcome.

The efficiency of direct transformation has been improved by the technique of electroporation; high field strength electrical pulses are used to reversibly permeabilise protoplast cell membranes to facilitate DNA uptake (Zimmermann and Vienken, 1982). This technique has been used to stably transform maize protoplasts (Fromm et al., 1986).

The overall view obtained from the studies on direct transformation is that transfer of DNA to most plant species is feasible: the only limitation of the methods discussed is the inability to regenerate transformed cells from all species into mature plants. An approach to resolving this problem is the use of viral vectors (for review see Hull and Davies, 1983) since plant viruses have the ability to spread

systemically through the plant, this would overcome the need for tissue culture and regeneration of the transformed plants. A severe limitation to the use of viruses as vectors is that they do not integrate into the plant genome to yield stably transformed plants. Cauliflower mosaic virus (CaMV) has been studied in detail in relation to its possible uses as a plant transformation vector. However, Gronenborn et al. (1981) have shown that the size of DNA insertions into the CaMV genome, due to viral packaging constraints, is limited to only ~250 bp; this is a severe limitation to its use as a vector. The modified virus is however still functional and will spread systemically through the plant but appears to be unstable over several rounds of growth. The work of Brisson et al. (1984) has shown that ORFII from CaMV can be replaced by a methotrexate-resistance gene, conferring resistance to the otherwise sensitive plant. The insertion size limit of this vector is however likely to be ~500 bp. The replacement of CaMV ORFII by the neomycin-phosphotransferase II gene (Paszkowski et al., 1986) has been shown to produce an avirulent CaMV molecule that integrates into the plant genome, an event not normally associated with CaMV infection. This potential vector however does not spread systemically through the plant and therefore does not overcome the limitations of tissue culture associated with some plant species. The modification of brome mosaic virus (BMV) to express chloramphenicol-acetlytransferase activity in barley protoplasts has been reported (French et al., 1985), this system is, however, still a long way from functioning as an effective vector system.

As a greater understanding of viral replicaion and

and infectivity is achieved. viral vectors may become available that will spread throughout the plant, carrying their cloned DNA with them, without causing the unwanted effects of viral disease.

One further approach to transforming whole plants unable to be regenerated from tissue culture was reported by Ohta (1986). This system relies on the introduction of DNA into maize embryos by pollination of maize flowers with a pollen/DNA mixture. The DNA, from a donor plant, contains dominant selectable markers that can be scored in the transgenic maize plants obtained from the transformed embryos. The exogenous DNA however appears to be unstable, giving low transformation frequencies. This paper is however flawed by the undefined nature of the DNA used for the transformation, and the lack of analysis of the fate of this transforming DNA.

A consideration of all the available methods for introducing and expressing foreign DNA in plant cells indicates that the Ti-plasmid, and more recently, direct transformation are at present the only reliable methods for introducing foreign DNA into plants.

4.E Expression of foreign genes in plants

The development of simple and rapid methods for the introduction of foreign genes into plant cells has led to the use of these methods in two distinct ways. The first is to investigate plant gene regulation; the second is for the modification of plants for agricultural purposes. An understanding of plant gene expression is however needed

before effective modification of plant genomes can be considered. The advances made towards understanding gene regulation by the use of Ti-plasmid-mediated plant transformation will be considered first, followed by some examples of plant modifications, of relevance to agriculture, that have been made.

The first attempts at expressing foreign genes, from bacteria, animals and unrelated plant species, in plant cells proved unsuccessful, as noted by Herrera-Estrella et al. (1983a) and Shaw et al. (1983). A solution to this problem was the use of the constitutive promoters from opine synthase genes to express the foreign genes. As discussed in Section I.4.B; this method of expressing foreign genes has proved to be very useful in providing selectable markers for transformed cells.

The first successful attempt at expressing a foreign gene in plant cells under its own promoter was the expression of bean phaseolin in sunflower callus (Murai et al., 1983). Under the control of its own promoter, this gene gave low levels of expression as compared to high levels when expressed by the nopaline synthase gene promoter. This work indicated that promoters from one plant species could function in an unrelated species and that tissue specificity plays an important role in gene-regulation. The expression of maize zein genes in sunflower callus tissue (Matzke et al., 1984; Goldsborough et al., 1986) indicated that promoters from monocotyledonous plants could function in cells from dicotyledonous plants. It was, however, unclear as to whether the low levels of expresion were due to the heterologous system or to the absence of the seed-specific

factors, necessary for high level zein gene expression, in the undifferentiated callus cells. Despite accurate transcription of the zein gene, the zein protein was not found in the callus tissue. More recently, following the development of vectors that enable transformed cells to be regenerated into whole plants, the organ-specific expression of foreign genes in whole plants has been described. Bean ß-phaseolin (Sengupta-Gopalan et al., 1985) and soybean ß-conglycinin genes (Beachy et al., 1985) have been transferred to tobacco and petunia plants respectively. These genes were correctly expressed in a seed-specific manner. A soybean leghaemoglobin gene, reported by Herrera-Estrella et al. (1983a) not to be expressed in transformed tobacco callus tissue, has been introduced into cells of the legume Lotus corniculata. Tissue-specific expression of the introduced gene in root-nodules was observed (Stougaard-Jensen et al., 1986).

The correct expression of heterologous heat-shock genes has been reported in transformed plants. Schoffl and Baumann (1985) have reported the expression of a soybean heat-shock gene (Hs) in heat-stressed transgenic sunflower tissue and Spena et al. (1985) have reported the heat-inducible expression of a chimaeric Drosophila Hs-notII gene in transformed tobacco plants. This observation, that the promoter from an animal gene can be activated by the appropriate signals in plant cells, indicates that there may be common gene-regulatory mechanisms for some plant and animal genes. The lack of expression of rabbit ß-globin in transgenic tobacco callus (Shaw et al., 1983) can possibly be explained by the absence of the correct tissue-specific

factors for gene induction in this tissue. The expression of a chimaeric nos-human growth hormone (HGH) gene in tobacco and sunflower callus tissue has demonstrated that these plant cells are unable to correctly process the HGH mRNA. This observation indicates a possible difference in the mRNA splicing apparatus between animal and plant cells. (Barta et al., 1986).

Along with research on tissue-specific expression of plant storage proteins, which may lead in the long term to the improved nutritional value of some crop plants, much attention has been directed towards the mechanisms of light-regulated gene expression; this may in turn lead to the increased productivity of crop plants. The use of Ti-plasmid vectors and plant transformation has provided much information on the mechanisms of light-regulated gene expression. The genes encoding the small subunit of Rubisco (Rubisco-SS) and the major chlorophyll a/b binding protein (AB) have been used extensively in these studies. The control of these genes was discussed in detail in Section I.2.C. The use of these two genes in chimaeric constructs with reporter genes, has enabled the identification and definition of the regions involved in light-regulation (see Section I.2.C). Jones et al. (1985) have shown that high levels of expression of genes introduced into homologous plants can be achieved. Broglie et al. (1984) and Nagy et al. (1985) have reported the expression of pea SS genes in transgenic petunia and tobacco. The levels of expression of this gene in transgenic petunia ranged from 0.2 to 10% of its level in pea; in tobacco, the average transcript level was 15x lower than petunia. The clonal variation between

plants is assumed to due to chromosomal position effects (Nagy et al., 1985). From these observations, it appears that genes introduced into heterologous plants are not expressed at the same high levels as they are when re-introduced into homologous plants.

The expression of foreign genes in tobacco has been used to study the mechanisms of protein transport to organelles. Van den Broeck et al. (1985) and Schreier et al. (1985) have both reported the transport of the bacterial protein nptII to the chloroplast by constructing chimaeric genes containing the transit peptide of the Rubisco-SS and the coding region of the nptII gene. Work by DeBlock et al. (1984) indicates that expression of foreign genes within the chloroplast is possible following selection for chloroplast transformation. Confirmation of chloroplast transformation is awaited with interest. These results indicate that not only the cytoplasm but the chloroplasts are amenable to modification by Ti-plasmid mediated transformation. Recent work by Simpson et al., (1986b) has shown, using chimaeric AB-nptII and SS-nptII genes, that expression of these genes is not only organ-specific in transgenic tobacco plants, but is cell-type-specific. These experiments have provided a greater understanding of plant cell biology, and a foundation for the modification of plant genomes for agricultural purposes. The development of a transient plant cell expression system (Werr and Lorz, 1986) will prove useful for the analysis of gene regulation and the role of tissue-specific factors in gene expression.

The use of Ti-plasmid transformation to introduce tobacco mosaic virus (TMV) coat protein genes into plants

(Bevan et al., 1986) will provide a means of understanding the processes involved in viral infection of plants. The expression of the TMV coat protein gene in transgenic tobacco has been shown to provide some cross-protection from subsequent TMV infection (Abel et al., 1986). The relevance of this observation to other plants and viruses is yet to be determined, but may result in major advances towards the protection of crop plants from viral disease.

A second area of commercial interest in plant transformation is in the area of herbicide-resistance. By engineering herbicide-resistance, or tolerance, into crop plants, the application of the herbicide to that crop will enable the selective destruction of weeds whilst preserving the crop. Two approaches have been used in order to develop glyphosate-tolerance in tobacco. Glyphosate is a herbicide that interferes with aromatic amino acid biosynthesis in the chloroplast. The first approach to engineering glyophosate-tolerance involves the isolation of a bacterial gene encoding a glyphosate-tolerant protein, and its transfer to plant cells under the control of a constitutive promoter (Comai et al., 1985). The second approach relies on the increased production of the wild-type herbicide-susceptible protein (Shah et al., 1985). Both approaches resulted in glyphosate-tolerance. Increased levels of tolerance could presumably be engineered by the transport, via a transit peptide, of the herbicide-tolerant proteins into the chloroplast. The herbicide, atrazine affects the Q_B protein which is part of the photosystem II complex. A herbicide resistant Q_B protein gene has been cloned (Golden and Haselkorn, 1985); the technology is

currently available for its expression in plant cells and its transfer to the chloroplast. Herbicide resistance can therefore be introduced into plants without recourse to plant breeding.

A major limitation to the manipulation of plant genomes at present is the inability of certain plant species to regenerate from tissue culture. Once this problem has been overcome, the considerable potential for the modification of crop plant can be realised.

5. AIMS OF THE PRESENT WORK

The availability of cDNA clones encoding the SS of Rubisco and the LHCP (Bedbrook *et al.*, 1980; S.M.Smith, unpublished results), provided a means by which genomic sequences for these genes could be isolated at the Warwick laboratory. At the time of starting this work, only one member of the SS multi-gene family from pea had been isolated and characterised (Cashmore, 1983); no pea AB genes had been isolated. The first aim of this project was thus the isolation and characterisation of other SS genomic sequences and AB genomic sequences from pea. During the course of this work, several SS genomic sequences were reported from pea and other plant species (see Section I.2.B). Two pea AB genomic sequences, and several from other plant species, have also been isolated (see Section I.3.B). The sequences presented in this thesis are however previously uncharacterised members of the SS and AB multi-gene families from pea.

Having isolated and characterised these genes, it was proposed that they be used to study the expression of the members of the two multi-gene families in both pea and transgenic tobacco plants. At the start of this work the vector technology had only just been developed to introduce foreign DNA into plant cells, and it had been used to investigate the expression of chimaeric genes under the control of T-DNA gene promoters; there had been no reports of the expression of plant genes under the control of their own promoters in transformed plant cells. During the course of this work however, several reports appeared of such expression (reviewed above).

The three basic aims of the work presented in this thesis were:-

1) To isolate and characterise genomic sequences encoding the small subunit of ribulose-1,5-bisphosphate carboxylase-oxygenase and the major chlorophyll a/b binding protein from pea (Pisum sativum).

2) To study the expression of these characterised genes, along with other members of the SS and AB multi-gene families, in pea plants.

3) To study the expression of the characterised SS and AB genes from pea in transgenic tobacco plants.

Much information regarding the SS and AB multi-gene families, plant transformation and the expression of foreign genes in plants, has been published during the course of this work. It is the author's belief, however, that the data presented in this thesis complement the results published recently by other researchers.

SECTION II

MATERIALS AND METHODS

1. MATERIALS

1.A Plant material

Seeds of Pisum sativum (var. Feltham First) were obtained from Charles Sharp, Boston Road, Sleaford, Lincs. Nicotiana tabacum (var. Samson) seeds were a generous gift from Dr.Michael Bevan of the Plant Breeding Institute, Cambridge.

1.B Chemicals, biochemicals, radiochemicals and enzymes

All materials used were of the highest analytical grade available. The source of specific reagents is given below.

Amersham International PLC, Amersham, Buckinghamshire: [α-^{32}P]dGTP (~3000Ci/mmol), [α-^{35}S]dATPαS (>400Ci/mmol), Hybond-N, DNA polymerase 1 (Kornberg), DNA polymerase 1 (Klenow fragment), restriction endonucleases, T4 DNA ligase, T4 polymerase.

The Boehringer Corporation (London) Ltd., Lewes, East Sussex: CIP,DNA polymerase 1 (Kornberg), DTT, Ficoll 400, proteinase K.

BDH Chemicals Ltd., Atherstone, Warwickshire: Acids, Amberlite monobed resin MB-3, AMPS, BPB, BME, m-cresol, HEPES, PEG 6000, 8-OH quinoline, phenol, solvents, TEMED, 1,1,1-trichloroethane, xylene cyanol.

Calbiochem-Behring, La Jolla, California,USA: NAA.

Difco Laboratories, Basingstoke, Hampshire: bacto-agar bacto-tryptone.

Eastman Kodak, Rochester, New York: Bis, TNS.

Fisons PLC, Loughborough, Leicestershire: acrylamide, CaCl, formamide, glycerol, SDS.

Flow Laboratories, Uxbridge Middlesex: MS plant salt mixture.

Gibco-BRL, Paisley, Renfrewshire, Scotland: BSA (nuclease free), lambda clind1ts857Sam7 DNA, LMP agarose (ultrapure), restriction endonucleases, urea (ultrapure).

Northumbria Biologicals Ltd., Cramlington, Northumberland: restriction endonucleases.

Oxoid Ltd., Basingstoke, Hampshire: nutrient broth, yeast extract.

New England Biolabs, Beverly, Massachusettes, USA: M13 pentadecamer sequencing primer.

Pharmacia (Great Britain) Ltd., Central Milton Keynes, Buckinghamshire: ddATP, ddCTP, ddGTP, ddTTP, δ-methacryloxypropyltrimethoxysilane.

Schleicher & Schull, Dassel, W-Germany: nitrocellulose type BA85/1.

Sigma Chemical Co. Ltd., Poole, Dorset: agarose medium EEO (type II), 4-amino salicylic acid, ampicillin (sodium salt), ATP, BSA, BA, carbenicillin (di-sodium salt), chloramphenicol, cysteine-HCl, dATP, dCTP, dGTP, dTTP, DMF, DNA (salmon sperm), DNAase 1 (bovine pancreas), EtBr, IPTG, kanamycin sulphate, N-lauroyl sarcosine (sodium salt), lysozyme, Mes, Mops, myo-inositol, nicotinic acid, nitrilo-triacetic acid, Pipes, poly(A)$^+$, pvp40, pyridoxine-HCl, RNAase A (bovine pancreas), S1 nuclease, spectinomycin 2-HCl, spermidine 3-HCl, tetracycline, thiamine-HCl, Trizma base, tRNA (E.coli strain W), Tween 20.

Worthington Diagnostic Systems Ltd., Freehold, New Jersey, USA: DNAase 1 (RNAase free).

All other chemicals were from either BDH Chemicals Ltd. or Fisons PLC depending on availability.

2. METHODS

2.A Growth and storage of biological materials

i) Storage of bacterial stocks

Bacterial stocks were stored as follows, (Maniatis et al., 1982).

For long term storage (over 9 months) bacteria were kept as frozen glycerol stocks:- 850 Ul of a 10 ml overnight culture in LB (10 g bactotryptone, 5 g yeast extract, 10 g NaCl, pH 7.5 per liter SDW) were mixed with 150 ul glycerol and stored at -70°C.

For medium term storage stocks were kept as stab cultures at room-temperature. Stab-agar (2 ml of 10 g Difco nutrient broth, 5 g NaCl, 8 g agar, 20 mg cysteine-HCl per liter SDW) in a 5 ml bijoux was inoculated from an overnight culture as above, incubated at 37°C overnight and stored at room temperature for between 6 and 9 months.

Short term storage of bacterial strains was on either minimal agar plates (5.25 g K_2HPO_4, 2.25 g KH_2PO_4, 0.5 g $(NH_4)_2SO_4$, 0.25 g tri-sodium citrate, 1.25 mg thiamine-HCl, 0.4 g glucose, 0.25 g $MgSO_4$ per liter SDW, solidified with bacto-agar to 1.5%), or LB agar plates (LB solidified with 1.5% bacto-agar). A loopful of a fresh overnight culture was spread onto a dried plate. Following incubation at 37°C overnight, plates were stored at 4°C for up to 2 weeks.

ii) Storage of phage stocks

a) Lambda phage

Lambda phage stocks were stored at 4°C in 5 ml phage buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$) containing 0.3% (v/v) chloroform (Maniatis et al., 1982).

b) M13 phage

M13 phage were stored either as frozen (-20°C) ssDNA in SDW or as RF DNA at 4°C in TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA)

iii) Growth of plants

Seeds of P.sativum (var. Feltham First) were sown in J.Arthur Bowers potting compost and grown at 20°C for 9 days in the light with a 12 h photoperiod using "warm-white" fluorescent lights (Philips). Light intensity was 50 μmol m^{-2} s^{-1}. Dark grown seedlings were grown in total darkness for 9 days.

Seeds of N.tabacum (var. Samson) were sown in the above potting compost and grown at 25°C in the light with an 18 h photoperiod at 100 μmol m^{-2} s^{-1}.

2.B Nucleic acid isolation

i) Large scale plasmid DNA extraction

Plasmid DNA was isolated from E.coli by a modification of the alkaline lysis method of Birnboim & Doly (1979). Ten ml of LB (10 g bacto-tryptone, 5 g yeast extract, 10 g NaCl, pH 7.5 per litre of distilled water) containing the appropriate antibiotic in 25 ml universal containers were inoculated with single bacterial colonies from fresh LB agar plates (as LB except solidified with Bacto-agar to 1.5%). Cultures were grown overnight in an orbital shaker at 200 rpm and 37°C. Two litre flasks containing 1 litre of LB plus antibiotics were inoculated with the 10 ml overnight cultures and grown at 37°C with shaking to an A_{590} value equal to 0.8 units. Plasmid DNA was amplified by the addition of either chloramphenicol to 175 µg/ml or, in the case of chloramphenicol-resistant plasmids, with spectinomycin to 300 µg/ml. The flasks were shaken at 37°C for a further 16 h. Cells were collected by centrifugation at 2500 rpm and 4°C for 30 min in an MSE 6x1000 rotor. The cells were resuspended in 3.2 ml of ice-cold GET (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8), transferred to 50 ml MSE Oakridge centrifuge tubes and placed on ice. Cell lysis was obtained by the addition of 200 µl GET containing lysozyme at 40 mg/ml and incubation on ice for 10 min; 6.6 ml of 0.2 M NaOH, 0.1% (w/v) SDS was added to each tube and the tubes gently swirled. After 10 min on ice, 5 ml of 3 M sodium acetate pH 4.5 was added and the contents of the tubes gently mixed. The tubes were left

on ice for a further 30 min. After this incubation the tubes were spun at 10,000 rpm and 4°C for 15 min in an MSE 8x50 rotor. Nucleic acid was precipitated by the addition of 8.3 ml isopropanol to the supernatants in sterile 50 ml tubes. Following mixing, the tubes were left at room temperature for 10 min. The precipitates were pelleted by centrifugation at 10,000 rpm and 4°C for 10 min. The pellets were drained and resuspended thoroughly in 3.2 ml 2 M ammonium acetate and the samples respun at 10,000 rpm for 10 min as above. After centrifugation, 2.1 ml isopropanol was added to each sample in sterile 50 ml tubes. After 10 min at room temperature the resulting precipitates were collected by centrifugation as above. The pellets were washed twice in 5 ml of 70% (v/v) aqueous ethanol. The plasmid pellets were gently dried *in vacuo* and resuspended in 20 ml TE (10 mM Tris-HCl pH 8, 1 mM EDTA).

Plasmid DNA was further purified by CsCl density gradient centrifugation (Maniatis *et al*., 1982). CsCl (23.76 g) was gently dissolved in each of the 20 ml plasmid solutions. EtBr (4 ml of 5 mg/ml) was added to each sample and the resulting solutions loaded into 37 ml Beckman Quickseal tubes. The tubes were topped up with liquid parafin, balanced and sealed. The gradients were formed by centrifugation in a Beckman Vti50 rotor run at 45,000 rpm and 20°C for 18 h. The rotor was brought to rest without the aid of the brake. Plasmid bands were visualised under u.v. light and withdrawn from the decapitated tubes in a volume of under 2 ml with an 18 gauge needle. The EtBr was removed from the samples by several extractions with isopropanol saturated with CsCl-saturated SDW. The CsCl was removed from

the sample by dialysis against several litres of TE. Dialysis tubing was prepared by boiling lengths of tubing in 0.25 M EDTA for 15 min followed by several rinses in sterile distilled water. After 3-4 h, the TE was changed and dialysis continued overnight. The plasmid DNA was stored at 4°C. A typical preparation yielded 500-1000 µg DNA.

ii) Small scale plasmid DNA extraction

Small scale plasmid DNA extractions were performed using a modified alkaline lysis protocol of Birnboim & Doly (1979) on either 2 ml overnight cultures in the case of plasmids in E.coli, or 10 ml 24 h cultures in the case of plasmids in A.tumefaciens. The cultures were grown in LB containing the appropriate antibiotic. The cells were collected by centrifugation and resuspended in 100 µl GET (50 mM glucose, 25 mM Tris-HCl pH 8, 10 mM EDTA). Resuspended cells were transferred to 1.5 ml Eppendorf centrifuge tubes and lysed by the addition of 20 µl GET containing lysozyme at 10 mg/ml and incubation on ice for 10 min; 200 µl of 0.2 M NaOH, 0.1% (w/v) SDS was added to the tubes and gently mixed. Following 5 min on ice 100 µl of 3 M sodium acetate pH 4.8 was added and incubation on ice continued for 10 min. Cell debris and chromosomal DNA was removed by centrifugation for 3 min. The supernatants were poured into new tubes and deproteinated by the addition of 200 µl TE saturated phenol and 300 µl chloroform/iso-amyl alcohol (24/1,v/v). The tubes were vortexed and spun in a microcentrifuge for 3 min. This step was repeated on the aqueous phase followed by extraction with 500 µl di-ethyl ether. RNA was removed from the samples by the addition of 5

µl boiled RNAase A (1 mg/ml). The plasmid DNA was precipitated by the addition of a half-volume of 7.5 M ammonium acetate, two volumes ethanol and incubation on dry-ice for 15 min. The DNA was collected by centrifugation at 4°C for 10 min in a microcentrifuge and the pellet washed twice in 70% (v/v) aqueous ethanol. The washed pellet was dried *in vacuo* and resuspended in 50 µl TE. In the case of plasmids isolated from *E.coli*, 5 µl of such a preparation was sufficient for analysis on an agarose gel but single-copy plasmids isolated from *A.tumefaciens* were obtained in such low amounts that the whole sample was needed for agarose gel analysis.

iii) Lambda phage DNA large scale extraction

Lambda phage stocks were prepared as follows, stored as described above (Section II.2.A.ii) and titred prior to DNA preparation (Maniatis *et al.*, 1982).

Plating cells were initially prepared. A single colony of *E. coli* strain K803 (Wood, 1966) from a fresh LB agar plate was inoculated into 10 ml LB (Section II.2.A.i) and grown at 37°C with shaking overnight. LB (80 ml) was inoculated with 1 ml of this overnight culture and incubated at 37°C with shaking for 3 h. The cells were harvested at 5,000 rpm and 4°C in MSE Oakridge centrifuge tubes in an 8x50 rotor. The cells were resuspended in 8 ml of 10 mM $MgSO_4$, shaken for 1 h at 37°C and stored at 4°C for up to 2 weeks.

Lambda phage stocks were prepared by taking 100 µl of plating cells and incubating them with $5x10^5$ pfu/ml of phage at 37°C for 20 min; these were then mixed with 3 ml

top-agar (LB supplemented with 10 mM $MgSO_4$ and solidified with 0.7% agar) at 45°C and plated onto undried LB plates supplemented with 10 mM $MgSO_4$. After overnight incubation at 37°C, 5 ml PB (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$) was added to the plates which were left at 4°C overnight. The top-agar and PB was scraped off the plates into sterile 30 ml Corex tubes (Du Pont). A further 1 ml PB was added to the plates; this was removed and added to the first 5 ml. Chloroform (0.18 ml) was added to the tubes which were vortexed. The top-agar and cell debris was pelleted by centrifugation at 10,000 rpm and 4°C for 10 min in an MSE 8x50 rotor. The supernatants were removed and placed in sterile 25 ml universal containers. Chloroform was added to 0.3% (v/v) and the stocks were vortexed then stored at 4°C.

Serial dilutions of the lambda phage sample to be titred were prepared in PB. Plating cells (100 µl) and 100 µl of each lambda dilution were mixed in 1.5 ml Eppendorf tubes and incubated at 37°C for 20 min. Each sample was mixed with 3 ml top-agar then quickly poured onto dried LB agar plates supplemented with 10 mM $MgSO_4$. Plates were incubated overnight at 37°C. A titre of pfu/ml of phage was calculated from the number of plaques per plate.

For large scale DNA preparations the method of Maniatis et al. (1982) was followed. A single colony of E.coli strain K803 was inoculated into 10 ml LB. The cultures were shaken at 37°C overnight, then used to inoculate 500 ml SB (35 g bacto-tryptone, 20 g yeast extract, 5 g NaCl, 2.5 g $MgSO_4$ per litre SDW pH 7) in 2 litre flasks. The bacteria were grown at 37°C and 250 rpm

until an A_{600} value of 0.3 was reached. For each lambda clone, 1.8×10^{10} pfu in 50 µl PB was added to the flasks of E.coli strain K803. The cultures were grown at 37°C and 250 rpm overnight. After at least 16 h the cultures were further lysed by the addition of 10 ml chloroform. The flasks were shaken as above for a further 30 min after which time the cultures had taken on a stringy appearance. Bacterial nucleic acid was removed by the addition of 100 µl DNAase 1 (1 mg/ml) and 50 µl RNAase A (5 mg/ml). After incubation at room temperature for 30 min, 30 g NaCl was added to each flask. The flasks were stood on ice for 1 h. Each 500 ml sample was divided equally between two 300 ml centrifuge bottles and spun for 20 min at 10,000 rpm and 4°C in an MSE 6x300 rotor. PEG 6000 (35 g) was dissolved in the supernatants from each pair of bottles and this mixture left at 4°C with occasional shaking for 1 h. The precipitate of phage was collected by centrifugation as above. The precipitates were drained well, resuspended in 8 ml PB, and transferred to 50 ml glass centrifuge tubes. The phage suspension was extracted once against an equal volume of chloroform. CsCl was added to the phage suspension to a final concentration of 0.5 g/ml. This suspension was layered onto a CsCl step gradient in 14 ml Beckman Ultra-clear centrifuge tubes. The steps were made in PB and consisted of, 1 ml *P*1.7 (1.27 g/ml) CsCl, 1.5 ml *P*1.5 (0.817 g/ml) CsCl and 1.5 ml *P*1.45 (0.706 g/ml) CsCl. The gradients were run in an SW40Ti rotor at 22,000 rpm and 4°C for 2.5 h. The phage band, which forms on the *P*1.45-1.5 interface, was removed using an 18 gauge needle. The phage sample was loaded into a 5 ml Beckman Ultra-clear centrifuge tube which

was topped up with 1.5 g/ml CsCl in PB. This second equilibrium gradient was loaded into an SW50.1 rotor and spun at 35,000 rpm and 4°C for 24 h. The band of phage was removed as above and dialysed for several hours at 4°C against several litres PB with at least two changes of this buffer. DNA was purified fron the phage particles by extraction twice with phenol and twice with chloroform/isoamyl alcohol (24/1,v/v). The DNA was precipitated by the addition of NaCl to 100 mM and two volumes of ethanol. The pellet was washed twice with 70% (v/v) aqueous ethanol, dried in vacuo. resuspended in TE and stored at 4°C.

iv) Lambda phage DNA small scale extraction

For small scale lambda DNA extractions a plate lysate method was used (Maniatis et al., 1982). Phage ($5x10^6$ pfu/ml) were incubated at 37°C with 100 µl of plating cells for 20 min. The mixture was added to 3 ml top-agarose (as top-agar except that agarose replaces bacto-agar) at 45°C and poured onto undried LB plates supplemented with 10 mM $MgSO_4$. The plates were incubated overnight at 37°C. PB (3 ml) was added to the plates which were placed at 4°C for 1 h. The PB and top-agarose were removed from the plates and placed in 50 ml MSE Oakridge centrifuge tubes. The tubes were spun at 10,000 rpm and 4°C for 10 min in an MSE 8x50 rotor. The supernatants were removed and placed at 4°C. A further 3 ml PB was added to each tube which was vortexed and centrifuged as before. The supernatants from the two spins were combined and centrifuged in 50 ml tubes at 14,000

rpm and 4°C for 20 min. An equal volume of 20% (w/v) PEG 6000, 2 M NaCl was added to the supernatants which were agitated and placed at 4°C for 1 h. The precipitates of phage were pelleted at 15,000 rpm and 4°C as above for 20 min. The pellets were thoroughly drained and resuspended in 500 µl PB with EDTA to 25 mM and SDS to 0.5% (w/v), then transferred to Eppendorf tubes. DNA was extracted as described above (Section II.2.B.iii).

v) M13 phage large scale replicative form DNA extraction

Competent cells of either E.coli strain JM101, (Messing et al., 1981), or JM103, (Messing et al., 1981), were transformed with either 20 ng M13 RF DNA or 1 µl ssDNA (Section II.2.B.vi). After incubation at 37°C overnight a single well-isolated plaque was inoculated into 1.5 ml 2xTY medium (16 g bacto-tryptone, 10 g yeast extract, 5 g NaCl per litre SDW) in a 25 ml universal container. The inoculum was grown at 37°C and 300 rpm for 6 h followed by storage at 4°C overnight. A 10 ml overnight culture of either E.coli strain JM101 or JM103 was grown in 10 ml 2xTY medium and used to inoculate 1 litre of 2xTY medium. The culture was grown to an A_{550} value of 0.5. To this culture was added the 1.5 ml sample that had been stored overnight and incubation was continued at 37°C and 300 rpm for 4 h. The RF DNA was isolated by the alkaline lysis method above described (Section II.2.B.i).

vi) M13 phage single stranded DNA extraction

The method used is from the M13 Cloning and Sequencing Handbook, Amersham (1984). A 1.5 ml aliquot of 2xTY medium containing a 1/100 dilution of an overnight culture of E.coli strain JM103 or JM101 was placed in a sterile 25 ml universal container which was inoculated with one plaque as described above (Section II.2.B.v). The culture was incubated at 37°C with vigorous shaking (300 rpm) for 5 h. The culture was transferred to a 1.5 ml Eppendorf tube and centrifuged in a microcentrifuge for 5 min. The supernatant was poured into a second tube and recentrifuged. The supernatant was again poured carefully into another tube and 200 µl of 20% (w/v) PEG 6000, 2.5 M NaCl added. Following mixing the tube was left to stand at room temperature for 15 min. The phage precipitate was pelleted by centrifugation in a microcentrifuge for 2 min. The supernatant was removed by pipette and the pellet respun after which any remaining PEG was removed. The viral pellet was resuspended in 100 µl TE and 50 µl of phenol saturated with TE buffer added. The tube was vortexed for 20 s followed by incubation at room temperature for 15 min. The sample was re-vortexed for 20 s and spun in a microcentrifuge for 5 min. The aqueous phase was re-extracted with 50 µl chloroform/iso-amyl alcohol (24/1,v/v). The ssDNA was precipitated after the addition of 10 µl of 3 M sodium acetate pH 6 and 250 µl ethanol at -20°C overnight. DNA was pelleted by centrifugation in a microcentrifuge at 4°C and the pellet washed twice in 70% (v/v) aqueous ethanol. The pellet was gently dried in vacuo.

resuspended in 30 µl SDW and stored at -20°C.

vii) Bacterial total DNA extraction

Bacterial total DNA was extracted as follows (Method from Dr.M.Bevan, personal communication). Nutrient broth (1.5 ml) containing kanamycin (50 µg/ml) was inoculated with an A.tumefaciens colony from a fresh agar plate and incubated at 30°C with shaking for 24 h. The culture was then transferred to a 1.5 ml Eppendorf tube and the bacteria collected by centrifugation. The pellet was resuspended in 300 µl 20 mM Tris-HCl pH 7.5, 5 mM EDTA; 100 µl 5% (w/v) N-lauroyl sarcosine (Na salt) was added followed by the addition of 50 µl TE containing proteinase K at 5 mg/ml. The sample was incubated at 37°C for 1 h after which it was passed twice through a 20 gauge needle. The material was extracted three times with a half-volume of TE-saturated phenol and a half-volume of chloroform/iso-amyl alcohol (24/1,v/v) followed by one extraction with di-ethyl ether. RNA was removed from the sample by the addition of 5 µl (1 mg/ml) boiled RNAase A and incubation at 37°C for 30 min. NaCl was added to a final concentration of 100 mM and two volumes of ethanol added. The DNA was precipitated by incubation on dry-ice for 15 min. The pellet was washed twice in 70% (v/v) aqueous ethanol, dried in vacuo and resuspended in 100 µl TE.

viii)Plant total DNA extraction

The following method (Dellaporta et al., 1983) was found to give high yields of high molecular weight genomic DNA. Leaf tissue (5 g) was frozen in liquid nitrogen and

ground to a fine powder with a pestle and mortar. To this powder was added 15 ml extraction buffer (100 mM Tris-HCl pH 8, 50 mM EDTA, 500 mM NaCl,10 mM BME) taking care not to allow the material to thaw until the buffer was added. This mixture was transferred to a 50 ml centrifuge tube and 1 ml of 20% (w/v) SDS was added and mixed thoroughly with the sample followed by incubation at 65°C for 10 min. Potassium acetate (5 ml, 5 M) was mixed vigorously with the sample and the whole incubated on ice for 20 min. The tube was spun at 4°C and 13,000 rpm in an MSE 8x50 rotor for 15 min. The supernatant was filtered through a Miracloth filter (Calbiochem) into a clean 30 ml Corex tube containing 10 ml isopropanol. The contents of the tube was mixed and incubated at -20°C for 30 min. The precipitated DNA was pelleted at 4°C and 10,000 rpm for 15 min in the above rotor and the pellet lightly dried by inverting the tubes over paper towels for 10 min. The pellet was redissolved in 700 µl 50 mM Tris-HCl, 10 mM EDTA pH 8 and transferred to a 1.5 ml Eppendorf tube. Any remaining insoluble material was removed by centrifugation for 10 min and the supernatant transferred to a clean tube. Sodium acetate pH 4.8 (75 µl, 3 M) followed by 500 µl isopropanol was added, mixed, and the clot of DNA pelleted for 5 min in a microcentrifuge. The pellet was washed twice with 70% (v/v) aqueous ethanol, dried, and allowed to redissolve overnight at 4°C in 100 µl TE. The DNA was stored at 4°C. RNA was removed from the sample prior to quantitation by the addition of 10 µl boiled RNAase A (1 mg/ml) and incubation at 37°C for 30 min.

ix) Plant total RNA extraction

The following protocol from McKnight & Palmiter (1979) as modified by Bennett et al. (1984), was used to extract total RNA from both pea and tobacco plants. Tissue that had been grown in the dark was harvested in the dark and frozen in liquid nitrogen prior to removal from the dark. All glassware was treated with chromic acid prior to use. Each frozen sample (1-2 g tissue) was homogenised in 10 ml phenol reagent (500 g redistilled phenol, 70 ml redistilled m-cresol, 0.58 g 8-OH quinoline saturated with 10mM Tris-HCl pH 7.6) and 10 ml Kirby reagent (6% (w/v) 4-amino salycylic acid, 10 mM Tris-HCl pH 7.6, 50 mM KCl, 1% (w/v) TNS, 9% (v/v) phenol reagent) using a Virtis homogeniser at half maximum speed for 30 s. The homogenate was transferred to a 30 ml Corex tube and centrifuged at 3,500 rpm and 4°C for 10 min in an MSE 8x50 rotor. The upper aqueous phase was removed and placed in a second 30 ml Corex tube. A half-volume of phenol reagent was added, and the sample mixed on a whirlimixer. A half-volume of chloroform was added, the sample was mixed again and then centrifuged as before. The aqueous layer (8 ml) was removed, and two volumes of ethanol added and then left at -20°C overnight. The RNA was pelleted by centrifugation as before, washed twice with 70% (v/v) aqueous ethanol containing 50 mM NaCl, drained and partly dried in vacuo. The pellet was resuspended in 2 ml of Mes-Mg reagent (30 mM Mes-NaOH pH 7, 20 mM magnesium acetate), transferred to a 15 ml Corex tube and 10 µl DNAase 1 stock (1 mg/ml RNAase free DNAase in Mes-Mg reagent) added. The sample was incubated on ice for 30 min followed by the addition of 1 ml phenol reagent. The

sample was mixed on a whirlimixer and centrifuged as before. The aqueous phase was re-extracted with a further half-volume of phenol reagent and a half-volume of chloroform. The aqueous phase was removed and sodium acetate added to a final concentration of 200 mM. Two volumes ethanol was added and the samples left overnight at -20°C. The RNA pellet was collected and washed as before, dried *in vacuo* and resuspended in 500 µl SDW. RNA was stored at -70°C

2.C Gel electrophoresis

i) High resolution agarose gels

When high resolution of DNA fragments was required agarose gels were made and run in 1x TAE buffer (10x TAE is 48.44 g Trizma base, 3.72 g EDTA per liter, pH 8.2 with acetic acid) as described by Maniatis et al. (1982). The agarose concentration used (0.6-0.8%) was dependent on the fragment size to be resolved (Maniatis et al., 1982). Agarose, medium EEO, type II was used except when the gel was a preparative one in which case LMP agarose was used. The appropriate weight of agarose was dissolved in 1x TAE using a microwave oven at low power for 5 min. The gel was cooled to 60°C before pouring. Leicester Biocenter gel tanks were used: 10 cm or 15 cm gels were poured using either 200 ml or 350 ml of 1x TAE agarose. The type of comb used was dependent on the sample size and the number of samples to be run on the gel. The gel was left to set for at least 1 h, then placed in the gel tank and covered in either 1.5 l or 2.5 l of 1x TAE buffer depending on the gel size. The samples were mixed with 5x loading buffer (50% glycerol, 50 mM EDTA, 0.1% bromophenol blue) and loaded into the wells; 10 cm gels were run at 35 V for 16 h and 15 cm gels at 60 V for 18 h. The gels were stained in 0.1 mg/ml EtBr for 30 min and were visualised on a u.v. transilluminator and photographed with a Polaroid instant camera using either Polaroid 665 or 667 film.

ii) Rapid analysis agarose gels

For rapid analysis of restriction enzyme digest

products or DNA quantification, a mini-gel tank (Uniscience) was used; 50 ml of 0.7% (w/v) agarose in 1x TBE buffer (10x TBE is Trizma base 108 g, boric acid 55 g, EDTA 9.5 g per liter DW) (Maniatis et al., 1982) buffer was prepared by melting the agarose in the buffer using a microwave oven. The molten gel was cooled to 60°C and EtBr added to 0.1 mg/ml. The gel was poured into the gel tank containing either an 8- or 16-place comb and left to set for 30 min at room temperature. The comb was removed and the gel covered in 1x TBE buffer; 5x loading buffer was added to each sample. The samples were loaded into the wells and the gel run for 30-60 min at 70 mA. Gels were visualised and photographed as described above (Section II.2.C.i)

iii) Denaturing polyacrylamide gels

a) Standard gels

The standard denaturing polyacrylamide gel was a 6% (w/v) polyacrylamide, 0.3% (w/v) bis gel in 7 M urea, 1x TBE (Section II.2.C.ii) (Amersham M13 Cloning and Sequencing Handbook, 1984). Gel plates (40 cm) were prepared by washing in detergent, rinsing in copious amounts of water and polishing with ethanol. When the gel was to be fixed to the back plate, this plate was treated with δ-methacryloxypropyltrimethoxysilane (Ansorge & De Mayer, 1980). If the gel was to be dried down onto paper this process was omitted. In all cases the notched plate was coated in 1,1,1-trichloroethane, allowed to dry and polished with ethanol. Plasticard gel combs (0.4 mm) were used and X-ray film spacers were used for sequencing gels whereas 0.4 mm plasticard combs and spacers were used for preparative

gels. The plates were taped together with vinyl tape. For a 50 ml gel 21 g urea (ultrapure) was dissolved in 7.5 ml 40% (w/v) acrylamide stock, 5 ml 10x TBE buffer, with distilled water to 50 ml. The 40% (w/v) acrylamide stock was made as follows: 38 g acrylamide, 2 g bis with distilled water to 100 ml. The solution was stirred with 5 g Amberlite MB-3 for 30 min and filtered through two layers of Whatman no.1 filter paper and stored in the dark. The 6% (w/v) acrylamide, 7 M urea, 1x TBE solution was filtered through a Millipore 0.45 μm filter and 300 μl freshly made 10% AMPS and 50 μl TEMED added. The solution was mixed and poured into the gel mould which was clamped with fold-back clips and the gel left for at least 1 h to polymerise. Gels were pre-run at 50 W in 1x TBE buffer for 30 min. and the slots rinsed out with 1x TBE buffer before the samples were loaded. Sequencing loading buffer (deionised formamide containing, 0.1% (w/v) xylene cyanol , 0.1% (w/v) bromophenol blue, 20 mM EDTA) was added to the samples which were then denatured by heating to 95°C for 3 min. Gels were run at 40 W for 1-3 h depending on the length of sequence to be read.

b)Buffer gradient gels

Gel plates (40 cm) were prepared as described above; 1x-2.5x buffer gradients were run to enable longer sequences to be read from a single gel (Biggin et al., 1983). The gels were made from 1x TBE and 2.5x TBE polyacrylamide stocks (as for 1x TBE polyacrylamide except that 2.5x TBE polyacrylamide contained 2.5x TBE and 10% [w/v] sucrose). Gels were run at 40 W with 0.5x TBE in the top tank and 1x TBE in the bottom tank for between 1.5-3 h.

2.D Nucleic acid restriction and modification reactions

i) Restriction endonuclease digestion of DNA

Restriction reactions were carried out according to the manufacturer's instructions with the following modifications.

Plasmid digestions were performed in reaction mixture volumes of between 20 and 100 µl with a DNA concentration below 100 ng/µl with 1-5 enzyme units/µg. Reaction times were between 1-2 h at 37°C exept in the cases of SmaI, when reactions were carried out at 30°C and TaqI, when the reaction temperature was 65°C. In the case of DNA that would not cut to completion under these conditions, spermidine was added to the reaction mixtures to a final volume of 4 mM.

Total bacterial DNA was treated in volumes of 100-200 µl with 1-5 enzyme units/µg for up to 12 h at a DNA concentration below 100 ng/µl. All such reaction mixtures contained spermidine at 4 mM

Digests of total plant DNA were performed in volumes of 100-200 µl for up to 3 h in the presence of 4 mM spermidine with between 5 and 10 units/µg.

Reactions were stopped either by the addition of gel loading buffer (Section II.2.C.i) in the case of samples to be analysed by agarose gel electrophoresis, or by phenol/chloroform extraction for samples that were to be used for further enzymatic manipulations.

Where appropriate, samples were concentrated following digestion by the addition of one half-volume of 7.5 M ammonium acetate, two volumes of ethanol followed by

incubation on dry-ice for 15 min. The pellet was washed twice in 70% (v/v) aqueous ethanol, dried and resuspended in the appropriate volume of either TE (Section II.2.B.i) or SDW.

ii) Dephosphorylation of vector DNA with calf intestinal phosphatase

Restricted plasmid DNA (1 µg) was extracted with a half-volume of phenol and a half volume of chloroform/iso-amyl alcohol (24/1,v/v), re-extracted with an equal volume of chloroform/iso-amyl alcohol, extracted with an equal volume of di-ethyl ether and precipitated by the addition of NaCl to 100 mM and two volumes of ethanol. After 15 min on dry-ice the DNA was pelleted by centrifugation and the pellet washed twice in 70% (v/v) aqueous ethanol, dried *in vacuo* and resuspended in 17 µl SDW. Removal of terminal 5'-phosphate groups was achieved as outlined in Maniatis *et al*. (1982); 2 µl of 10x CIP buffer (500 mM Tris-HCl pH 9, 10 mM spermidine, 10 mM $MgCl_2$, 1 mM $ZnCl_2$) was added along with 1 µl of 1 unit/µl CIP. For 3'-recessed ends the reaction mixture was incubated at 37°C for 30 min followed by the addition of a further unit of CIP and a further 30 min incubation. In the case of 5'-recessed or blunt ends the 30 min incubations were replaced by the following two incubations. 15 min at 37°C followed by 15 min at 56°C. The reaction was terminated by the addition of nitrilo-triacetic acid to 10 mM and heating to 70°C for 10 min to denature the CIP. The vector was then precipitated and washed as above (Section II.2.B.ii) and resuspended in 10 µl SDW. Dephosphorylated vector DNA was stored at 4°C.

iii) Ligation of DNA with T4 DNA ligase

Ligation reactions were carried out in 10 µl volumes containing 50-100 ng of dephosphorylated vector and 50-100 ng of the insert fragment. T4 DNA ligase (1 unit) was used per ligation and 1 µl of fresh 10x ligation buffer (0.5 M Tris-HCl pH 7.5, 0.1 M $MgCl_2$, 0.05 M DTT, 0.01 M HCC, 0.01 M spermidine, 0.01 M ATP and 1 µg/µl BSA [nuclease-free]) (modified from Rusche & Howard-Flanders, 1985) added. The reactions were incubated at 15°C for 16 h and 5 µl used per bacterial transformation.

2.E Isolation of DNA following electrophoresis

i) Elution of DNA from agarose gels

This method is a modification of that used by Dretzen et al. (1981) and was found to give very high yields of DNA that was essentially free from any inhibitors of DNA restriction and modification enzymes. DNA fragments to be isolated were subjected to electrophoresis through 1x TAE LMP agarose gels (Section II.2.C.i). A slot was cut in front of the desired band and a piece of Whatman DE81 paper (pre-soaked overnight in 2.5 M NaCl,20 mM Tris-HCl pH 7.5, 1 mM EDTA, rinsed 5 times in SDW and stored at 4°C in 1 mM EDTA) inserted into the slot. The gel was re-run until the DNA band had bound to the paper. The paper was removed from the gel and placed in 300 µl per 50mm^2 NTE buffer (1.5 M NaCl,20 mM Tris-HCl pH 7.5, 1 mM EDTA) in a 1.5 ml Eppendorf tube, vortexed and incubated at 37°C for 2 h followed by a further 10 min at 50°C. The paper was pelleted by centrifugation at room temperature for 10 min. The supernatant was then spun-filtered through two layers of Whatman no.1 paper into a clean Eppendorf tube. The yield of DNA was increased by spinning a further 50 µl NTE buffer through the filter paper. The EtBr was removed from the sample by extraction three times with an equal volume of water-saturated butanol. The DNA was concentrated by the addition of two volumes ethanol and incubation on dry-ice for 15 min. The precipitated DNA was pelleted by centrifugation, washed twice with 70% (v/v) aqueous ethanol, dried, and resuspended in SDW to a concentration of 50 ng/µl.

ii) Elution of DNA from polyacrylamide gels

For the isolation of small [^{32}P]-labelled DNA fragments from 1x TBE, standard polyacrylamide gels (Section II.2.C.iii.a) the band was first located by exposure to Fuji RX X-ray film for 5-15 min. The band was excised and placed in 300 µl PAGEB (0.5 M ammonium acetate, 0.01 M magnesium acetate, 0.1% SDS (w/v), 0.1 mM EDTA), modified from Maniatis et al. (1982) in a 1.5 ml Eppendorf tube and incubated at 37°C overnight. The polyacrylamide gel slice was pelleted by centrifugation and 10 µg E.coli tRNA added to the supernatant. DNA was precipitated by the addition of three volumes ethanol followed by incubation on dry-ice for 15 min. The precipitate was pelleted by centrifugation, washed twice with 70% (v/v) aqueous ethanol, dried, and resuspended in a small volume of SDW.

iii) Electro-elution of DNA from polyacrylamide gels

In cases where DNA fragments were to be eluted rapidly the fragment was located as described above and the excised gel slice placed in 400 µl 1x TBE buffer (Section II.2.C.ii). The buffer and gel slice were placed in EDTA-treated dialysis tubing (Section II.2.B.i). The slice was placed against one edge of the tubing which was sealed at both ends with dialysis tubing clips. This tubing was subjected to electrophoresis in 1x TBE buffer for 30 min at 150 V in a mini-gel tank; the polarity was reversed for 10 s and the buffer removed to a 1.5 ml Eppendorf tube. The gel slice was rinsed with 100 µl 1x TBE and this buffer added to the original 400 µl. The DNA was precipitated and purified as described above (Section II.2.E.ii).

2.F Radio-labelling of DNA

i) Nick-translation of DNA with DNA polymerase 1

The methods employed were modified from Rigby et al. (1977). High specific activity labelled DNA was used for hybridisation probes whereas low specific activity labelled DNA was used for restriction mapping work.

a) High specific activity labelling

Where the DNA was to be labelled to a high specific activity the nick-translation reactions were carried out in a volume of 30 µl in a 0.7 ml Eppendorf centrifuge tube; 1 µl of 100 ng/µl DNA to be labelled was mixed with 3 µl 1x nick-translation buffer (10x buffer consists of 0.5 M Tris-HCl pH 7.8, 50 mM $MgCl_2$, 100 mM DTT), 1 µl each of 100 µM dATP, dCTP, dTTP, 1 µl of 1 µM dGTP, 3 µl of 50 ng/ml DNAase 1 and 1 µl [α-^{32}P]dGTP. The volume was made up to 29 µl with SDW and the reaction mixture incubated at 37°C for 15 min; 1 µl of DNA polymerase 1 was added to the tube and the reaction further incubated for 3 h at 15°C. The specific activity of DNA labelled in this way was typically $1x10^8$ dpm/µg. After the incubation 5 µg E.coli tRNA carrier was added to the labelled DNA and the unincorporated nucleotides removed by two sequential ammonium acetate/ethanol precipitations (Maniatis et al., 1982). The pellet was washed twice in 70% (v/v) aqueous ethanol, dried and resuspended in 100 µl SDW.

b) Low specific activity labelling

Where DNA fragments were to be labelled for restriction mapping very low specific activities were

required. To achieve this, 100 ng DNA was labelled as described above with the following exceptions; DNAase 1 was omitted from the reaction mixture, only 1 µl (10 µCi) of [α-^{32}P]dGTP was added, and the reaction was carried out at ambient temperature for 30 min. The reaction was stopped by heating to 65°C for 10 min and the unincorporated nucleotides removed as described above.

ii) End-labelling with:

a) Klenow polymerase

This method was employed for end-labelling DNA restriction fragments with 3'-recessed ends. For the preparation of labelled DNA markers for agarose gels, 1 µg restricted DNA was ethanol-precipitated, washed twice with 70% (v/v) aqueous ethanol, dried and resuspended in 15 µl SDW. To this solution was added 1 µl nucleotide mix containing 2 mM dATP, 2 mM dCTP and 2 mM dTTP and 1 µl (10 µCi) of [α-^{32}P]dGTP. Following the addition of 1 µl (2 u) of Klenow polymerase the reaction mixture was incubated at ambient temperature for 30 min. In the case of DNA to be labelled for restriction mapping the above protocol was followed with the exeption that 100 ng DNA was used instead of 1 µg. Reactions were terminated by heating to 65°C for 10 min and unincorporated nucleotides removed as for nick-translated DNA.

b) T4 DNA polymerase

Where DNA restriction fragments with 5'-recessed ends were to be end-labelled for restriction mapping the following protocol was used. DNA (100 µg) in 10 µl SDW was mixed with 1.5 µl 10x nick-translation buffer (Section

II.2.F.i.a) and 1 µl (1 u) of T4 DNA polymerase. The reaction mixture was incubated at 37°C for 3 min after which time 1 µl nucleotide mix (Section II.2.F.ii.a) was added along with 1 µl (10 µCi) of [α-^{32}P]dGTP and the reaction continued at 37°C for 40 min. The reaction was terminated and the unincorporated nucleotides removed as described in Section II.2.F.i.a.

iii) Uniform labelling of single-stranded DNA with Klenow polymerase

The method of Burke (1984) was followed. Single-stranded DNA (250 ng) and sequencing primer (50 ng) was annealed in a volume of 242 µl in prime-cut probe buffer (20 mM Tris-HCl pH 7.5, 2 mM DTT, 12 mM $MgCl_2$) by heating to 90°C for 3 min and cooling to 65°C for 10 min. The solution was quenched on ice and dATP, dCTP and dTTP added to 100 µM. The annealed template was stored at -70°C. To prepare sufficient probe for S1 analysis on 10 samples, 50 µl of the annealed template was used. The template was thawed and 30 µCi of [α-^{32}P]dGTP added. The polymerisation reaction was started by the addition of 1 µl (5 u) of Klenow polymerase. The reaction mixture was incubated at 18°C for 20 min followed by the addition of dGTP to 1 mM and a further incubation at 18°C for 20 min. The reaction was terminated by heating the sample to 65°C for 15 min. The reaction was cooled to 4°C and made 50 mM with respect to NaCl. Restriction endonuclease HindIII (20 u) was added and digestion allowed to proceed at 37°C for 30 min. The sample was precipitated by the addition of 10 µg of E.coli tRNA and two volumes of ethanol. The pellet was washed twice in 70%

(v/v) aqueous ethanol. dried and resuspended in 10 µl SDW; 4 µl of sequencing sample buffer (Section II.2.C.iii.a) was added and the sample denatured at 95°C for 3 min. The sample was loaded onto a standard 6% (w/v) denaturing polyacrylamide gel (Section II.2.C.iii.a) and run for 1 h at 45 W. The gel was exposed to Fuji RX X-ray film with registration marks for 5 min. After location of the band on the gel it was excised and electro-eluted from the polyacrylamide (Section II.2.E.iii). and resuspended in 50 µl SDW.

2.G Nucleic acid hybridisation

i) Southern hybridisation

a) Cloned DNA

For the analysis of DNA cloned into lambda phage or plasmid vectors the following modified procedure of Southern (1975) was followed. DNA to be analysed was subjected to electrophoresis through 1x TAE agarose gels (Section II.2.C.i). After visualisation of the DNA bands, the gels were soaked twice in twice their volume of 1.5 M NaCl, 0.5 M NaOH for 20 min. The gels were neutralised by soaking twice in twice their volume of 1.5 M NaCl,0.5 M Tris-HCl pH 7.4 for 20 min. Nitrocellulose was prepared by soaking for 30 min in SDW followed by 30 min in 20x SSC (3 M NaCl, 0.3 M tri-sodium citrate pH 7.0). The DNA was transferred onto the nitrocellulose by capillary blotting in 20x SSC for 2-24 h. Following transfer the filter was baked *in vacuo* at 80°C for 2 h. The filter was pre-hybridised in a sealed bag for 1 h at 42°C in 200 µl pre-hybridisation buffer per cm^2 of filter. Pre-hybridisation buffer consists of:- 1x SSC, 5x Denhardt's solution (100x Denhardt's solution is 2% pvp40, 2% Ficoll 400 and 2% BSA) and 50% formamide, (deionised by stirring with amberlite MB-3 monobed resin), and sheared denatured salmon sperm DNA to 200 µg/ml (Maniatis *et al.*, 1982). Following pre-hybridisation the filter was sealed in a fresh bag containing 200 µl hybridisation buffer per cm^2 of filter. Hybridisation buffer consists of:- 1x SSC, 1x Denhardt's solution, 50% formamide pH 7.0, sheared denatured salmon sperm DNA to 200 µg/ml (Maniatis *et al.*, 1982). Typically 10^6 dpm of nick-translated probe (Section

II.2.F.i.a) was used per hybridisation. Hybridisation was carried out at 42°C for 16-24 h. Filters were washed at room temperature three times for 15 min each time in 2x SSC, 0.1% SDS and three times at 60°C for 15 min each time in 0.1x SSC, 0.1% SDS. The filters were dried at room temperature and autoradiographed with a "Lightening Plus" intensifier screen (Cronex) at -70°C for 1-3days.

b) Genomic DNA

Where genomic DNA was to be analysed the above protocol was modified as follows (Amersham Membrane Transfer and Detection Methods, 1984). Prior to denaturation of the DNA in the agarose gel it was depurinated by soaking in twice the gel volume of 0.25 M HCl for 15 min. The DNA was denatured and neutralised as described above, and the DNA transferred by capillary blotting onto Amersham Hybond-N membrane. Following transfer the membrane was rinsed in 2x SSC, wrapped in cling-film, and exposed, DNA side down, on a u.v. transilluminator for 2-5 min. The membrane was pre-hybridised for 2 h at 65°C in 200 µl per cm^2 hybridisation buffer (6x SSC, 5x Denhardt's, 0.5% (w/v) SDS) containing 100 µg/ml poly(A)$^+$. Hybridisation was performed at 65°C for 16 h in 200 µl per cm^2 hybridisation buffer containing 100 µg/ml poly(A)$^+$ and $1x10^7$ dpm/µg denatured nick-translated probe (Section II.2.F.i.a). The membranes were washed twice at room temperature in 2x SSC, 0.1% (v/v) SDS for 15 min and once at 55°C in 0.2x SDS, 0.1% (v/v) SDS. Membranes were sealed in bags and exposed to Kodak X-Omat S X-ray film with an intensifying screen for 2-10 days.

ii) DNA dot-blot hybridisation

DNA dot-blot hybridisations were carried out according to Corruzi et al. (1984). DNA (100 ng) at 20 ng/µl was denatured by the addition of 5 µl 0.6 M NaOH. After 20 min the solution was neutralised by the addition of 40 µl 0.6 M Tris-HCl pH 3. The DNA samples were loaded onto nitrocellulose that had been soaked for 2 min in SDW followed by 20 min in 0.6 M Tris-HCl pH 7 on a dot-blot apparatus. The nitrocellulose was allowed to dry and baked at 80°C in vacuo. Pre-hybridisation and hybridisation was performed in 50% formamide, 0.6 M NaCl, 50mM $NaPO_4$ pH 7, 5x Denhardt's containing 100 µg/ml single-strand denatured salmon sperm DNA. The hybridisation contained $3x10^4$ dpm of nick-translated probe (Section II.2.F.i.a). After 16 h hybridisation, the filters were washed at the appropriate stringency, dried and exposed to Kodak X-Omat S X-ray film with an intensifing screen for 24 h.

iii) Colony hybridisation

Bacterial colonies carrying recombinant plasmids that were to be screened by colony hybridization as described by Grunstein & Hogness (1975) were streaked onto a nitrocellulose filter disc placed on an LB agar antibiotic plate. The plates were grown overnight at 37°C. The disc was then removed and placed colony-side up on 4 layers of Whatman 3MM paper soaked in 10% (v/v) SDS. After 3 min the filters were removed and placed on another 4 layers of Whatman 3MM soaked in 0.5 M NaOH, 1.5 M NaCl for 5 min. The next stage was repeated twice; the filters were placed on 4 layers of Whatman 3MM paper soaked in 0.5 M Tris-HCl pH 8,

1.5 M NaCl for 5 min. The discs were finally placed on 4 layers of Whatman 3MM soaked in 2x SSC and after 5 min blotted dry between a few layers of Whatman no.1 paper. After 30-60 min, drying the filters were baked *in vacuo* for 2 h at 80°C. Prehybridisation, hybridisation, washing and autoradiography were as described in (Section II.2.G.i.a)

iv) Plaque hybridisation

Plaques of either lambda or M13 phage, to be screened by hybridisation following the method of Benton & Davis (1977), were transferred to nitrocellulose discs by placing the discs onto fresh plaque-containing agar plates that had been pre-chilled to 4°C for 30 min. The discs were left on the plates for 1 min. The phage DNA was released from the phage particles and denatured by placing the disc on 4 layers of Whatman 3MM paper soaked in 0.5 M NaOH, 1.5 M NaCl for 1 min. The filters were neutralised for 5 min on Whatman 3MM soaked in 0.5 M Tris-HCl pH 8, 1.5 M NaCl and transferred for 5 min to 3MM paper soaked in 2x SSC. The filters were air-dried, baked, pre-hybridized, hybridized, and autoradiographed as described in (Section II.2.G.i.a).

2.H Bacterial transformation

i) Transformation of E.coli with plasmid DNA

The following method was obtained from Lamb (1984). An overnight culture of the appropriate E.coli strain was established in 10 ml Psi broth (2% (w/v) bacto-tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl and 20 mM $MgCl_2$). This overnight culture (0.1 ml) was used to inoculate a further 10 ml of warm Psi broth and this culture grown at 37°C with shaking to an A_{550} value of 0.3. This culture (0.5 ml) was inoculated into 25 ml of warm Psi broth which was grown as above to an A_{550} of 0.48. The cells were harvested at 4°C and 5,000 rpm for 5 min in an MSE 8x50 rotor. The cells were gently resuspended in 10 ml ice-cold Tfb no.1 (100 mM RbCl, 50 mM $MnCl_2$, 10 mM $CaCl_2$, 35 mM sodium acetate pH 5.8 and 15% glycerol) and incubated on ice for 15 min. The cells were harvested as above and gently resuspended in 1 ml ice-cold Tfb no.2 (10 mM RbCl, 75 mM $CaCl_2$, 10 mM Mops pH 5.8 and 15% glycerol). The cells were incubated on ice for 1-24 h. The volume of competent cells used per transformation was 100 µl. The DNA, up to a volume of 10 µl, was added to the cells and these were left on ice for 30 min. The cells were then heat-shocked at 42°C for 2 min. The next stage was dependent on the type of DNA used in the transformation. When the DNA was plasmid DNA, and the selection was antibiotic-based, the heat-shocked cells were allowed to recover prior to plating by the addition of 1 ml Psi broth and incubation without shaking for 1 h at 37°C. The cells were spun down in a microcentrifuge for 2 min,

resuspended in 100 µl Psi broth and plated on dried LB antibiotic plates. Antibiotics were used at the following concentrations for plasmid selection:- ampicillin 50 µg/ml; chloramphenicol 10 µg/ml; kanamycin 25 or 50 µg/ml (see Section II.2.I.i); tetracycline 25 µg/ml. In the case of plasmid DNA where selection was based on the ß-galactosidase expression system, the heat-shocked cells were added to 3 ml LB, 0.7% agar at 42°C containing 25 µl of 25 mg/ml X-Gal in DMF and 25 µl of 25 mg/ml IPTG. The mixture was vortexed and poured onto dried LB agar plates. When M13 DNA was used in the transformation, the heat-shocked cells were added to 3 ml HTOP (10 g bacto-tryptone, 8 g NaCl, 8 g agar per liter SDW) at 42°C to which had also been added 100 µl of a 3 h culture of the E.coli strain used in the transformation and 25 µl of 25 mg/ml X-Gal and 25 µl of 25 mg/ml IPTG. The mixture was then poured onto minimal agar plates (Section II.2.A.i). Plates were incubated at 37°C overnight. Competent cells prepared in this way gave a transformation efficiency of about $1x10^7$ colonies or plaques per µg of DNA.

2.I Plant transformation and tissue culture

i) A.tumefaciens-E.coli mating

This procedure is from Bevan (1984). Overnight cultures of E.coli strain MC1022 (Casadaban & Cohen, 1980) harbouring Bin19 (Bevan, 1984) constructs, and E.coli strain HB101 (Boyer & Roulland-Dussoix, 1969) harbouring pRK2013 (Figurski and Helinski, 1979), were grown separately in 10 ml LB containing kanamycin (25 µg/ml) at 37°C with shaking. A culture of A.tumefaciens strain LBA4404 (Hoekema et al., 1983) was also grown overnight at 30°C in 10 ml NB containing streptomycin (500 µg/ml). Samples (200 µl) of each of the three overnight cultures were mixed and streaked onto a dried LB-agar plate. The plate was incubated overnight at 30°C. A loopful of the culture from this overnight plate was streaked onto an agrobacteria minimal media agar plate (210 g K_2HPO_4, 90 g KH_2PO_4, 4 g $MgSO_4$, 0.2 g $CaCl_2$, 0.1 g $FeSO_4$, 0.04 g $MnCl_2$, 20 g NH_4Cl and 0.5% glucose per liter SDW solidified with 1.5% agar) containing kanamycin (50 µg/ml). The plate was incubated at 30°C for 3 days. A single colony was restreaked twice more on the above plates to give a stock of A.tumefaciens ready for plant transformation.

ii) Plant transformation

The methods employed were as outlined by Horsch et al. (1985). A 24 h culture of A.tumefaciens was grown at 30°C in 10 ml NB containing kanamycin (50 µg/ml). The cells

were pelleted by centrifugation at 4°C and 5,000 rpm for 5 min in an MSE 8x50 rotor. The cells were resuspended in 10 ml NB containing no antibiotics. Young leaves from approximately 70-day old plants of N.tabacum var. Samson were surface-sterilised for 30 min in 10% (v/v) Domestos containing a few drops of Tween 20 per liter SDW. The leaves were rinsed several times in SDW under sterile conditions and 6 mm diameter discs cut from the leaves using a cork-borer. The discs were floated in the resuspended culture in a 5 cm diameter Petri dish sealed with Nesco film. The dish was incubated at room temperature for 3 h with gentle shaking. After incubation with the agrobacteria,the leaf discs were blotted dry and placed on nurse culture plates (Fraley et al., 1983. Nurse culture plates were prepared as follows: a 7 cm filter paper was placed on a shoot-induction plate (4.6 g MS salts, 1x B5 vitamins [100x B5 vitamins contains:- 5 g myo-inositol, 0.5 g thiamine-HCl, 0.05 g nicotinic acid, 0.05 g pyridoxine-HCl], 3% (w/v) sucrose, 0.1 mg NAA, 1 mg BA, per liter SDW pH 5.7, solidified with 0.7% agarose) and either tobacco or potato suspension cells (0.5ml - a gift from Dr.M.Bevan) layered onto the filter. A second 5 cm filter disk was placed over the suspension cells to complete the nurse culture plates. Nurse culture plates with leaf-discs were sealed with Nesco film and placed in a growth room at 25°C under 18 h, 100 μmol $m^{-2}s^{-1}$ illumination for 3 days. After 3 days the leaf-discs were transferred to shoot-induction plates; these were made as described for nurse culture plates except that the filter papers and suspension cells were omitted. Shoot-induction plates

contained kanamycin (300 μg/ml) and carbenicillin (500 μg/ml). The cultures were grown under the above conditions; callus tissue that grew was subcultured onto fresh plates every 3 weeks. As shoots appeared over the next 4-6 weeks they were excised and transfered onto root-induction plates (made as for shoot-induction plates but with the omission of phytohormones) containing kanamycin (100 μg/ml) and carbenicillin (500 μg/ml). Roots appeared over the following 1-2 weeks and the plantlets were transferred to soil as soon as possible after this. After transfer to soil the plants were covered with clear polythene bags to maintain high humidity until they gained control of their water balance which was usually one week.

iii) Surface-sterilisation of tobacco seeds

Seeds that were to be germinated under antibiotic-selection were surface-sterilised in 20% (v/v) Domestos for 40 min followed a rinse in 70% (v/v) aqueous ethanol and several rinses in SDW. The seeds were plated out individually onto root-induction plates containing kanamycin (100 μg/ml) and germinated under the same conditions as outlined in Section II.2.A.iii.

2.J S1 nuclease analysis

i) RNA-probe annealing and S1 nuclease digestion

The following method was modified from Burke (1984). RNA (1,5 or 10 µg) was ethanol-precipitated, washed twice in 70% (v/v) aqueous ethanol, lightly dried and resuspended in 5 µl SDW. Uniformly labelled single-stranded DNA probe (5 µl of 20 pg/µl) (Section II.2.F.iii) was added to the RNA. The samples were mixed and sealed in capillary tubes. Samples were denatured by heating to 95°C for 5 min and annealing carried out for 3 h at the appropriate temperature. The optimum temperature for maximal annealing was ascertained by annealing at temperatures with 5°C intervals between 50°C and 70°C. After annealing the reaction mixtures were quenched on ice and the contents of each capillary transferred to 150 µl ice-cold S1 buffer (0.28 M NaCl, 0.05 M sodium acetate pH 4.6, 4.5 mM $ZnSO_4$, 20 µg/ml sonicated, denatured, salmon sperm DNA) containing 100 units S1 nuclease. The reaction mixtures were incubated at 30 °C for 30 min followed by the addition of 3 µl 0.4 M EDTA. The samples were extracted once with 150 µl TE-saturated phenol and the nucleic acid precipitated by the addition of 340 µl ethanol. The pellets were washed twice in 70% (v/v) aqueous ethanol, dried and resuspended in 3 µl sequencing loading buffer (Section II.2.C.iii.a). Samples were denatured at 95°C for 3 min and run on a standard 6% (w/v) denaturing polyacrylamide gel (Section II.2.C.iii.a) for 1 h at 45 W. Gels were exposed to Fuji RX X-ray film at -70°C overnight.

2.K DNA sequencing

DNA sequencing was performed using the dideoxy chain termination method of Sanger et al. (1977) using either [α-^{32}P]dGTP as the radio-labelled nucleotide, or [α-^{35}S]dATPαS (Biggin et al., 1983). The method described in the Amersham M13 Cloning and Sequencing Handbook (1984) was followed.

i) Sequencing with [^{32}P]

For each of the four bases, deoxynucleotide zero-mixes were made as outlined in Table 2. Equal volumes of zero mixes and ddNTP solutions were mixed using 0.5 mM ddTTP, 0.2 mM ddCTP, 0.125 mM ddGTP and 0.2 mM ddATP. Aliquots (2 µl) of each of these zero-mix/dideoxynucleotide solutions were placed in 1.5 ml Eppendorf tubes on ice. Single-stranded template DNA (5 µl), derived as described in Section II.2.B.vi, was annealed to 1 µl of sequencing primer in Klenow reaction buffer (10 mM Tris-HCl pH 8, 5 mM $MgCl_2$) in a final volume of 10 µl for 2 h at 60°C in an oven. The annealed template was cooled on ice and 0.5 µl (5 µCi) [α-^{32}]PdGTP added. Klenow polymerase was diluted to 1 u/µl in SDW and 1 µl added to the annealed template. Samples of template/label/enzyme mix (2.5 µl) were placed onto the rim of the four 1.5 ml Eppendorf reaction tubes. The reactions were started by a flick-spin in a microcentrifuge and incubated at ambient temperature for 15 min. After the initial reaction time, 1 µl of 0.5 mM dGTP chase solution was added to the reaction mixtures and incubation continued for a further 15 min. The reactions were terminated by the addition of 4 µl of sequencing loading buffer (Section

Table 2. Composition of nucleotide zero-mixes used for sequencing with [α-^{32}P]dGTP

The nucleotide zero-mixes used for sequencing when the labelled nucleotide was [α-^{32}P]dGTP (Section II.2.K.i) were made using the indicated amounts of 0.5 mM dNTPs and TE buffer.

	T°	C°	G°	A°
0.5 mM dTTP	1 µl	20 µl	20 µl	20 µl
0.5 mM dCTP	20 µl	1 µl	20 µl	20 µl
0.5 mM dGTP	1 µl	1 µl	1 µl	1 µl
0.5 mM dATP	20 µl	20 µl	20 µl	1 µl
TE buffer	5 µl	5 µl	5 µl	5 µl

Table 3. Composition of nucleotide zero-mixes used for sequencing with [α-^{35}S]αSdATP

The nucleotide zero-mixes used for sequencing when the radio-labelled nucleotide was [α-^{35}S]αSdATP (Section II.2.K.ii) were made using the indicated amounts of 0.5 mM dNTPs and TE buffer.

	T°	C°	G°	A°
0.5 mM dTTP	1 µl	20 µl	20 µl	20 µl
0.5 mM dCTP	20 µl	1 µl	20 µl	20 µl
0.5 mM dGTP	20 µl	20 µl	1 µl	20 µl
TE buffer	20 µl	20 µl	20 µl	20 µl

II.2.C.iii.a) and heating to 95°C for 15 min. Samples (1 µl) were loaded onto either standard 6% (w/v) polyacrylamide gels or buffer-gradient gels (Sections II.2.C.iii.a & b). Short-run standard gels were run until the BPB dye front was 5 cm from the bottom of the gel and long-run standard gels were run until the xylene cyanol dye front was 5 cm from the bottom of the gel. Buffer gradient gels were run slightly longer in both cases. Gels were fixed in 10% (v/v) acetic acid, 10% (v/v) methanol and dried either onto the back-plate in an 80°C oven, or onto Whatman 3MM paper. Sequencing gels were exposed to Fuji RX X-ray film at room temperture for 12-36 h.

ii) Sequencing with [^{35}S]

When sequencing was performed using [^{35}S] the above annealing procedure (Section II.2.K.i) was followed and the deoxynucleotide zero mixes made according to Table 3. Equal volumes of the zero-mixes and dideoxynucleotide solutions were mixed using 0.5 mM ddTTP, 0.02 mM ddCTP, 0.05 mM ddGTP and 0.15 mM ddATP and placed in tubes as described above (Section II.2.K.i). The sequencing procedure was followed as in Section II.2.K.i with the following differences. The initial reaction time was increased from 15 min to 20 min. 1 µl of [α-^{35}S]dATPαS (>400Ci/mmol) added to the template/label/ enzyme mix instead of [α-^{32}P]dGTP and the reaction was chased with a uniform mix of deoxynucleotides at 0.5 mM.

iii) Clone selection by T-tracking

To differentiate between several randomly cloned fragments in M13 phage, T-tracking was employed. The method outlined below is a modification from the Amersham M13 Cloning and Sequencing Handbook (1984). The T zero-mix/ddTTP solution was made as in Table 1 and Section II.2.K.i. The annealing reaction was mixed as follows: 0.5 µl single-stranded DNA was mixed with 0.5 µl sequencing primer and 2 µl of 1x Klenow reaction buffer (Section II.2.K.i), and annealing performed by heating the samples to 95°C for 5 min and cooling the reaction to room temperature over 15 min. Zero-mix/ddTTP solution (2 µl) was added to the annealed template and these mixed by a flick-spin in a microcentrifuge. For 30 samples 3 µl (30 µCi) of [α-^{32}P]dGTP was diluted in 15 µl SDW containing 7.5 u Klenow polymerase. The enzyme/label mix (0.5 µl) was added to the annealed DNA and the reaction incubated at room temperature for 15 min. The reaction mixture was chased, stopped, denatured, and subjected to electrophoresis as described in Section II.2.K.i.

SECTION III

RESULTS AND DISCUSSION

1. ANALYSIS OF LAMBDA GENOMIC CLONES

1.A Introduction

The genomic clones characterised in this work were isolated from two λEMBL3 (Frischauf *et al*., 1983) genomic libraries constructed and screened entirely by Dr. T. F. Gallagher. The DNA used for the construction of both libraries was isolated from *P. sativum* (var. Feltham First).

The lambda clones containing SS genomic sequences were isolated from a library containing total *Bgl*II digested genomic DNA. The library was screened with the combined cDNA inserts from the plasmids pSSU60 and pSSU160 (Bedbrook *et al*., 1980). Eight lambda clones showing homology to the cDNA inserts were isolated from this library and these were designated as follows; λSS47, λSS48, λSS49, λSS51, λSS52, λSS54, λSS56 and λSS64.

The second library was constructed from genomic DNA that had been partially digested with *Mbo*I. This library was screened with the cDNA clone pF*a*/*b*31 (Dr.S.M.Smith unpublished results). This clone contains a partial AB cDNA sequence. The first lambda clone isolated showing homology to pF*a*/*b*31 was designated λAB13. All lambda genomic clones identified as containing SS or AB sequences were screened for homology to pBR322 and all found to show none.

1.B Identification of the genomic equivalent of SSU60

One of the requirements of this project was the isolation of an expressed member of the SS gene family. It has been shown in pea (Coruzzi *et al*., 1984) as well as in

other species (Broglie et al., 1983; Dean et al., 1985b) that members of the SS multi-gene family show sequence divergence in their 3'-untranslated regions. These regions can therefore be used as gene-specific probes to isolate the genomic equivalent of a cDNA as shown by Coruzzi et al. (1984). A comparison of the 3'-untranslated regions of all known pea SS genes is shown in Figure 1. Since the SSU60 cDNA was reverse-transcribed from a polyadenylated message the genomic equivalent of SSU60 must be transcriptionally active. It therefore follows that isolation of the genomic equivalent of SSU60 would result in the isolation of an expressed member of the SS multi-gene family. The isolation of the genomic equivalent of SSU60 would also lead to the analysis of an, as yet, uncharacterised member of the SS multi-gene family from pea.

The 3'-untranslated region of SSU60 used as a gene-specific probe is shown in Figure 2. This sequence has been shown to differ from that published by Bedbrook et al. (1980) (Dr.S.M.Smith, personal communication). It is presumed that the sequence published by Bedbrook et al. (1980) is that of SSU1. Plasmid DNA was prepared as described in Section II.2.B.i. The cDNA insert was isolated from pSSU60 by cleavage with HindIII (Section II.2.D.i) and isolated from a 1% LMP agarose gel (Section II.2.E.i). The purified insert was further cleaved with HinfI to yield a 74 bp fragment from the structural region of the gene, and a fragment ~170 bp long containing 8 bp of structural sequence, 148 bp of 3'-untranslated sequence, plus the poly(A^+) tail. The larger of the two fragments was isolated from a 1.2% LMP agarose gel (Section II.2.E.i). As a

FIGURE 1. Sequence comparison between the 3'-untranslated region of SSU60 and other members of the SS multi-gene family from pea

The upper sequence shown is the gene-specific probe from SSU60. The sequence starts at the HinfI site and ends at the polyadenylation site. The corresponding regions from other reported pea SS genes, SS3.8 (Cashmore, 1983), PS2.4 (Coruzzi et al., 1984), SS8.0 (Timko et al., 1985a) and SSU1 (Bedbrook et al., 1980), are aligned with SSU60 according to the above sites. The coding region of these sequences is underlined. Base substitutions between SSU60 and the other genes are indicated by an asterisk (*) whereas deleted bases are indicated by a dash (-). The percentage homology between SSU60 and the other sequences is indicated by the figure in brackets preceding the sequence. The sequences are arranged in order of homology to SSU60. The length of the sequence between the conserved HinfI site and the polyadenylation site is indicated by the figure in brackets following the sequence. The vertical arrows indicate the deduced S1 cleavage point (based on a 2 bp mismatch) for a RNA-DNA hybrid between the genomic equivalent of SSU60 (SS47) and the RNA corresponding to these other genes. (See Section III.4.A).

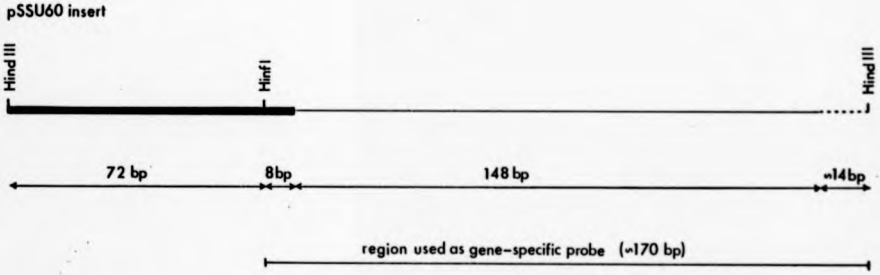

pSSU60 insert
Hind III
Hinf I
Hind III
72 bp
8bp
148 bp
≈14bp
region used as gene-specific probe (≈170 bp)

FIGURE 2. Isolation of a gene-specific probe from SSU60

The cDNA insert from pSSU60 is shown. The thick line represents the coding region, the thin line represents the 3'-untranslated region and the dotted line represents the poly(A^+) tail. The insert was cleaved from the plasmid with HindIII and isolated from an agarose gel. The cDNA was then cleaved with HinfI and the larger (~170 bp) fragment isolated from on agarose gel. This HindIII-HinfI fragment, which contains 8 bp of coding region, 148 bp of 3'-untranslated region and ~14 bp of poly (A^+) tail, was used as a gene-specific probe to isolate the genomic equivalent of SSU60.

positive control the cDNA insert was cleaved from pSSU160 with HindIII and isolated from a 1% LMP agarose gel. This insert contains only coding sequence; the coding region of SS genes has been shown to be highly conserved, within species, between different members of the multi-gene family (Coruzzi et al., 1984; Dunsmuir et al., 1983a). These inserts were independently nick-translated with DNA polymerase 1 (Section II.2.F.i.a), and used to probe replicate dot-blot filters of all eight λSS clones (Section II.2.G.ii). λDNA was prepared as described in Section II.2.B.iv. The replicate filters were washed twice at 37°C for 20 min in 0.1% SDS and either 2x SSC, 1x SSC or 0.5x SSC. The filters were dried and autoradiographed overnight. The results are shown in Figure 3. From these data it was inferred that λSS47 contained the genomic sequence (SS47) most closely related to SSU60.

It was later shown by sequence analysis (Section III.3.B) that SS47 is in fact the genomic equivalent of SSU60. A comparison of the 3'-untranslated region of SSU60 (Dr.S.M.Smith, personal communication) and SS47 (Section III.3.B) showed 100% sequence homology between these two sequences. A comparison of the 3'-untranslated region of SSU60 with the corresponding region from the published SS sequences from pea showed homologies ranging from 67-89% (Figure 2). The genes compared with SS47 in Figure 1 were all isolated from P.sativum var. Progress No.9. It was initially thought that intervarietal differences would prevent meaningful comparison between genes from P.sativum var. Feltham First and var. Progress No.9; as shown in Section III.3 and Section III.4 this is not the case.

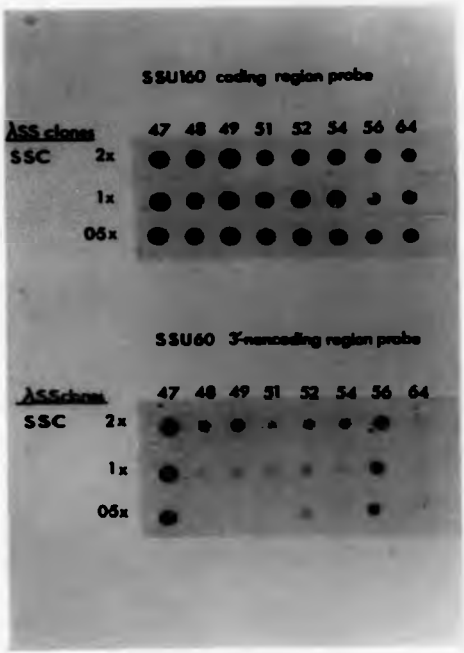

SSU16O coding region probe
ΔSS clones
47 48 49 51 52 54 56 64
SSC 2x
1x
05x
SSU60 3'-noncoding region probe
ΔSS clones
47 48 49 51 52 54 56 64
SSC 2x
1x
05x

FIGURE 3. Hybridisation of pSSU60 to eight small subunit lambda genomic clones

A. DNA (100 ng) from eight lambda genomic clones shown to carry SS genes was loaded onto nitrocellulose using a dot-blot apparatus. The designated numbers of these clones is indicated above the dots. Replicate dots were probed with nick-translated cDNA insert from pSSU160 and washed at 37°C at the indicated SSC concentration twice for 15 min. Filters were exposed to Kodak X-Omat overnight with an intensifying screen.

B. Replicate dot blots (as described above) were probed with the gene-specific probe described in Figure 1. The filters were washed at the indicated SSC concentration and autoradiographed as described above.

Intervarietal comparisons can be made since the same members of the SS multigene family appear to be present in both varieties of pea. From the data presented above, it is concluded that SS47 is an expressed member of the SS multi-gene family in P.sativum.

1.C Restriction and Southern analysis of lambda SS47

Having identified λSS47 as containing an expressed SS gene it was further characterised by restriction analysis with several 6 bp recognition-site restriction endonucleases and Southern hybridisation. The purpose of this was to identify restriction fragments that could be subcloned into plasmid vectors from where SS47 could be characterised in greater detail.

Lambda DNA was prepared as described in Section II.2.B.iii. The restriction digests were performed as outlined in Section II.2.D.i and the cleavage products separated on a 0.6% agarose gel (Section II.2.C.i) using λHindIII digestion products as size markers. The gel was stained with EtBr and photographed (Figure 4). The DNA was transferred to nitrocellulose (Section II.2.G.i.a) and probed with nick-translated cDNA insert from pSSU160. The filters were washed (Section II.2.G.i.a) and exposed to Fuji RX X-ray film overnight. The autoradiograph is shown in Figure 5. From the agarose gel and Southern blot, restriction fragments containing SS47 were identified.

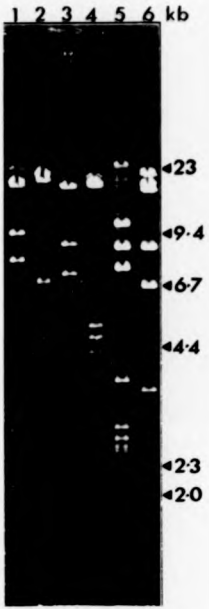

1 2 3 4 5 6 kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 4. Restriction endonuclease analysis of λSS47

λSS47 DNA (1 μg) was digested with BamHI, lane1; BglII, lane 2; EcoRI, lane 3; HindIII, lane 4; PstI, lane 5 and SalI, lane 6, and the digestion products analysed on a 1x TAE agarose gel run at 1.5 V/cm for 16 h. The sizes indicated in kb are those of HindIII cleaved lambda DNA. The gel was stained in 0.1 g/ml EtBr for 15 min following electrophoresis.

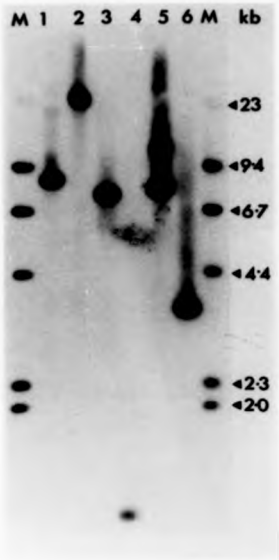

M 1 2 3 4 5 6 M kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 5. Southern analysis of restriction endonuclease digested λSS47 DNA

DNA restriction fragments were transferred from the agarose gel described in Figure 4 to nitrocellulose by capillary blotting. The lanes are as labelled in Figure 4. The filter was probed with nick-translated cDNA insert from pSSU160 for 16 h. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat with an intensifying screen. HindIII cleaved λDNA, end-labelled with [α-^{32}P]dGTP, was used as size markers; the sizes of the fragments are shown in kb (lane M). The 23 kb markers appear faint, possibly due to inefficient transfer to the nitrocellulose

1.D Restriction and southern analysis of lambda AB13

The identification of an expressed AB gene was not possible using the technique employed for the isolation of SS47. The cDNA pFa/b31 is not fully characterised; it was therefore not possible to isolate a gene-specific probe from this clone. λAB13 was the first pFa/b31 homologous clone isolated from the library. For this reason, and the inability to positively identify any uncharacterised AB gene isolated as an expressed member of the multi-gene family, λAB13 was chosen for further characterisation. The expression of AB13 is investigated in Section III.4. As with λSS47, the first stage in this characterisation was the analysis of the products of 6 bp recognition-site restriction endonucleases. The restriction digests and agarose gel electrophoresis were performed as for λSS47 (Section III.1.C). The results of this analysis are shown in Figure 6. The DNA was transferred to nitrocellulose and probed with nick-translated pFa/b31 insert as described in Section II.2.F.i.a. The blot was washed as outlined in Section II.2.G.i.a and autoradiographed overnight. The results are shown in Figure 7. From these data it was possible to identify restriction fragments containing AB13 that could be subcloned into plasmid vectors.

1.E Discussion

From Figure 3A it can be seen that when the eight lambda SS clones are probed with a DNA fragment from the coding region of the gene (SSU160) there is virtually no reduction in the amount of probe hybridised to each clone as the stringency of the wash is increased. This indicates the


1 2 3 4 5 6 kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 6. Restriction endonuclease analysis of λAB13

λAB13 DNA (1 μg) was digested with BamHI, lane1; BglII, lane 2; EcoRI, lane 3; HindIII, lane 4; PstI, lane 5 and SalI, lane 6, and the digestion products analysed on a 1x TAE agarose gel run at 1.5 V/cm for 16 h. The sizes indicated in kb are those of HindIII cleaved lambda DNA. The gel was stained in 0.1 g/ml EtBr for 15 min following electrophoresis.

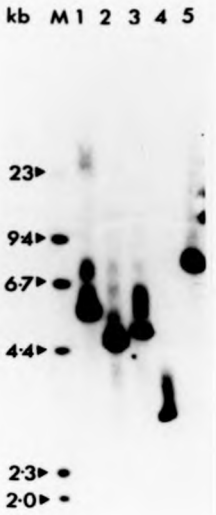

kb M 1 2 3 4 5
23
9·4
6·7
4·4
2·3
2·0

FIGURE 7. Southern analysis of restriction endonuclease digested λAB13 DNA

DNA restriction fragments were transferred from the agarose gel described in Figure 6 to nitrocellulose by capillary blotting. The lanes are as labelled in Figure 6 with the exeption of lane 6 which did not transfer to the nitrocellulose and is therefore absent. The filter was probed with nick-translated cDNA insert from pF*a*/*b*31 for 16 h. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat with an intensifying screen. The smears present above the hybridising bands are artifacts and do not correspond to restriction fragments (see Figure 6). HindIII cleaved λDNA, end-labelled with [α-^{32}P]dGTP was used as size markers; the sizes of the fragments are shown in kb (lane M).

high degree of homology between the probe and filter-bound DNA. Comparing the levels of hybridisation to each clone, it can be seen that λSS49 shows a stronger hybridisation signal than any other clone. As the coding region of SS genes has been shown to be highly conserved between members of the multi-gene family (Coruzzi et al., 1984 and Dunsmuir et al., 1983a) it is likely that this particular clone contains more than one SS gene. It was been noted by Cashmore (1983) that SS8.0 and SS3.6 were isolated from the one genomic fragment cloned into lambda. Dean et al (1985c) also noted that petunia SS genes are clustered within the genome and one lambda clone may contain more than one SS genomic sequence. The levels of hybridisation to λSS56 and λSS64 appear to be lower than for the other clones. This could be due to either these lambda clones containing only part of a SS sequence or the sequences contained by these clones being pseudogenes and divergent from SSU160. In Figure 3B the same eight lambda clones give very different hybridisation patterns when probed with the 3'-untranslated region from SSU60. Only λSS47 retains the same level of hybridisation when probed with the 3'-specific sequence as it did when probed with the coding region. All other clones show far weaker homology to the 3'-specific probe even at the lowest stringency washes. As the washing stringency is increased the level of hybridisation to λSS47 remains similar. The level of hybridisation of the 3'-specific probe to all the lambda clones other than λSS47 drops dramatically as the washing stringency is increased.

It can therefore be concluded that of these eight lambda clones λSS47 shows the greatest sequence homology to

the 3'-untranslated region of SSU60. λSS56 and λSS52 show some weak homology to SSU60, whereas the other lambda clones appear even more divergent in their 3'-untranslated region. From these data and a comparison of the sequence data presented in Figure 1 with that of SSU60, it is concluded that SS47 is the genomic equivalent of SSU60 and that it must therefore be an expressed member of the multi-gene family. By comparing the homology of the 3'-untranslated region of SSU60 to the equivalent region of the other SS sequences known, namely SSU1 (Bedbrook et al., 1980), PS2.4 (Corruzi et al., 1984), SS3.6 (Cashmore 1983), and SS8.0 (Timko et al. 1985a), it can be seen that SS3.6 shows the highest homology within this region followed by PS2.4 (Figure 1). The SS gene that shows the least homology to SSU60 is SSU1. It is not possible to say conclusively which of the lambda clones examined in Figure 3 contain these already characterised sequences. The reason is that there may be as yet uncharacterised members of the SS multi-gene family (Section III.4.B).

From the restriction enzyme and Southern hybridisation data of SS47 it can be seen that more than one BglII fragment has been cloned into the vector (Figure 4, lane 2). The fragment containing the gene is ligated to the vector left arm; the two extra fragments are both ~3 kb. The library screened for SS sequences was constructed by the ligation of BglII fragments into BamHI-cleaved vector; neither the BglII, nor the BamHI site is reconstituted by this ligation. Any BglII sites in the lambda clone will therefore be due to either sites in the vector, or sites in the insert and not sites at the vector/insert ligation

point. The vector arms will generate a 20 kb fragment (the left arm) and a 7.3 kb fragment (the right arm). Several other small fragments of between 400 bp and 600 bp will be generated from the right arm. The extra BglII sites in the insert can be explained in the following ways; firstly, the genomic digest may not have gone to completion and the BglII fragments ligated into the vector are contiguous. The second possibility is that the multiple inserts in the vector are not contiguous in the pea genome. The first possibility does not constitute a problem as far as isolating a complete gene is concerned. The second possibility will only cause a problem when the extra restriction site is internal to the region of interest. Later work (Section III.3.B) shows that there are no BglII sites within the SS gene or its immediate flanking regions. From the Southern blot data presented in Figure 5, it can be seen that the fragments containing SS sequences are an 8.2 kb BamHI fragment; a BglII fragment >23 kb (this fragment contains the vector left arm and part of the insert); a 7.2 kb EcoRI fragment; a 1 kb HindIII fragment; a 7.7 kb PstI fragment. The second PstI fragment, ~10 kb, that hybridises to the SS cDNA probe is deduced to be a partial digestion product composed of the 7.7 kb fragment and a fragment of about 2.5 kb. This is concluded from the facts that λSS47 contains only one SS gene; other genes would show extra hybridisation bands with other enzyme digests, and there are no internal PstI sites in or in close proximity to the SS47 gene (Section III.3.B). The final band of interest is the 3.6 kb SalI fragment. As mentioned above, it can be concluded that λSS47 contains only one SS gene. This gene is located on a BglII fragment of ~10 kb and can

be isolated from the lambda clone on any one of the fragments identified in Figure 5.

Similar conclusions can be drawn for λAB13 from Figure 6 as for λSS47 from Figure 4. The insert in the vector is ~16 kb. Data presented in Figure 7 identify the fragments that carry AB13 as: a 5.4 kb BamHI fragment, with some homology to the vector left arm: a 4.5 kb BglII fragment; a 4.9 kb EcoRI fragment; a 3.0 kb HindIII fragment and a 7.7 kb PstI fragment. The SalI restriction fragments shown in Figure 6, lane 6 did not transfer well onto the nitrocellulose and this lane is not shown in Figure 7. As found with λSS47, λAB13 contains only one gene. This was concluded from the presence of single hybridisation fragments to the restriction products of all the enzymes used.

In conclusion it can be stated that λSS47 is the genomic equivalent of SSU60, an as yet uncharacterised member of the SS multi-gene family. The genomic sequence, SS47, can be isolated from the lambda clone by several 6 bp recognition site restriction endonucleases for subcloning into plasmid vectors. λAB13 contains an AB gene, the expression of which is investigated in Section III.4. Several restriction endonuclease-generated fragments from λAB13 containing the AB gene have been identified and can be isolated for subcloning into plasmid vectors for characterisation in greater detail. The sizes of the restriction fragments containing both SS47 and AB13 are different to any fragments previously shown to contain members of these multi-gene families (Cashmore, 1983; Cashmore, 1984; Coruzzi *et al.*, 1984; Timko *et al.*,

1985a). Due to the uncertainty of whether the BglII inserts in λSS47 are contiguous in the genome or not, the restriction fragment sizes may not represent the fragment size found in the genome. The restriction mapping and sequence analysis of these genes (Sections III.2 and III.3) confirms their previously uncharacterised status.

2. SUBCLONING OF RESTRICTION FRAGMENTS FROM LAMBDA CLONES INTO PLASMID VECTORS

2.A Strategy

The restriction fragments from λSS47 and λAB13 identified in Section III.1.D as containing the genomic sequences SS47 and AB13 respectively were subcloned into plasmid vectors to enable characterisation of these sequences in greater detail. Plasmid pAT153 (Twigg & Sherratt, 1980) was used as a recipient for BamHI and HindIII restriction fragments whereas plasmid pACYC184 (Chang & Cohen, 1978) was used as a recipient for EcoRI fragments. Plasmid DNA was prepared by the method given in Section II.2.B.ii, cleaved with the appropriate restriction enzyme (Section II.2.D.i) and dephosphorylated as outlined in Section II.2.D.ii. Lambda DNA was prepared as described in Section II.2.B.iii. Lambda DNA was restricted with the appropriate restriction enzyme and the digest products shotgun-ligated into the corresponding vector (Section II.2.D.iii). E.coli strain DHI (Hanahan, 1983) was transformed with the ligation products (Section II.2.H.1). Colonies containing plasmids were selected by growth on either ampicillin (pAT153 ligations) or tetracycline (pACYC184 ligations).

Plasmids containing either SS47 or AB13 sequences were distinguished from both religated vector, and vector containing other sequences, by colony hybridisation as outlined in Section II.2.G.iii. Colonies derived from ligations with λSS47 DNA were probed with nick-translated

cDNA insert (Section II.2.F.i.a) from pSSU16O; colonies originating from ligations with λAB13 DNA were probed with nick-translated cDNA insert from pFa/b31 (data not shown). Colonies that gave positive signals were further analysed by restriction enzyme digestion of small scale plasmid DNA preparations (Section II.2.B.ii). The restriction digest products were run on an agarose gel (Section II.2.C.i) with linearised plasmid and λHindIII cleavage products as markers (data not shown).

Plasmids containing the desired inserts were prepared as outlined in Section II.2.B.i. The inserts were further verified by electrophoresis of plasmid and lambda clone DNA cleaved with the same enzymes (Figures 8 & 10). These gels were Southern-blotted and probed with either pSSU16O insert DNA or pFa/b31 insert DNA as appropriate (Figures 9 & 11). Plasmids produced were designated; pSS47B (containing an 8.3 kb BamHI restriction fragment derived from λSS47), pSS47E (containing a 7.4 kb EcoRI restriction fragment derived from λSS47), pSS47H (containing a 1.0 kb HindIII restriction fragment derived from λSS47), pAB13B (containing a 5.4 kb BamHI restriction fragment derived from λAB13) and pAB13E (containing a 4.9 kb EcoRI restriction fragment derived from λAB13). Having confirmed that the sequences cloned into the plasmid vectors were indeed either SS47 or AB13 they were subjected to restriction enzyme analysis and restriction maps produced.


1 2 3 4 5 6 kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 8. Restriction endonuclease analysis of SS47 subclones

Plasmid DNA (1-3 μg) of the three λSS47 subclones, pSS47B, pSS47E and pSS47H was cleaved with BamHI. lane 2; EcoRI. lane 4 and HindIII, lane 6 respectivly and electrophoresed on a 1x TAE agarose gel at 1.5 V/cm for 16 h. The plasmid DNA was run alongside λSS47 DNA (1 μg) cleaved with BamHI. lane 1; EcoRI. lane 3 and HindIII. lane 5. The size markers in kb indicate the positions of the restriction products from λDNA cleaved with HindIII. The gel was stained in 0.1 g/ml EtBr for 15 min following electrophoresis.

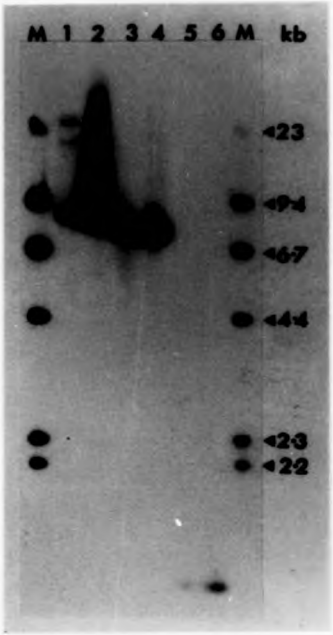

M 1 2 3 4 5 6 M kb
23
9·4
6·7
4·4
2·3
2·2

FIGURE 9. Southern analysis of SS47 plasmid subclones

The DNA fragments fractionated on the gel described in Figure 8 were transferred to nitrocellulose by capillary blotting. The filter was probed with nick-translated insert from pSSU160 for 18 h. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat with an intensifying screen. HindIII cleaved λDNA, end-labelled with [α-^{32}P]dGTP was used as size markers; the sizes of the fragments are shown in kb (lane M). The smear in lane 2 is due to sample overloading (see Figure 8, lane 2).


1 2 3 4 kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 10. Restriction endonuclease analysis of AB13 subclones

Plasmid DNA (1 μg) of the two λAB13 subclones pAB13B and AB13E was cleaved with BamHI, lane 2 and EcoRI, lane 4 respectively and electrophoresed on a 1x TAE agarose gel at 1.5 V/cm for 16 h. The plasmid DNA was run alongside λAB13 DNA (1 μg) cleaved with BamHI, lane 1 and EcoRI, lane 3. The size markers in kb indicate the positions of the restriction products from λDNA cleaved with HindIII. The gel was stained in 0.1 g/ml EtBr for 15 min following electrophoresis.


1 2 3 4 M kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 11. Southern analysis of AB13 plasmid subclones

The DNA fragments fractionated on the gel described in Figure 10 were transferred to nitrocellulose by capillary blotting. The filter was probed with nick-translated insert from pSSU160 for 16 h. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat with an intensifying screen. HindIII cleaved λDNA, end-labelled with [α-^{32}P]dGTP was used as size markers; the sizes of the fragments are shown in kb (lane M).

2.B Restriction enzyme analysis of plasmid subclones containing SS47

The method employed for producing restriction maps of the SS47 subclones involved the isolation of the plasmid insert (Section II.2.E.i). Purified insert (100 ng) was radio-labelled by nick-translation (Section II.2.F.i.b), and 100 ng end-labelled with either Klenow polymerase or T4 DNA polymerase (Sections II.2.F.ii.a & b). Labelled plasmid inserts (10 ng) were subjected to restriction endonuclease digestion with various enzymes (typically 0.5-1 unit). The restriction fragments were analysed by agarose gel electrophoresis. The gels were dried down and autoradiographed overnight. End-labelled fragments were run alongside nick-translated fragments digested with the same enzyme; this enabled the identification of the terminal fragments produced during digestion. An example of such a gel is shown in Figure 12. Various double digests were performed and an outline of a restriction map produced.

The positions of restriction sites were, where necessary, confirmed by EtBr-stained agarose gel analysis of the plasmid clone. Plasmid DNA (1 µg) was digested as outlined in Section II.2.D.i and analysed on a 0.7% agarose gel (Section II.2.C.i). An example of an EtBr-stained agarose mapping gel is shown in Figure 13. The enzymes used in the construction of the maps were as follows:- AvaI. BamHI. BglII. EcoRI. HindIII. NdeI. PstI. SalI. SmaI. SphI and XhoI. The restriction maps of pSS47B, pSS47E and pSS47H are shown in Figure 14.

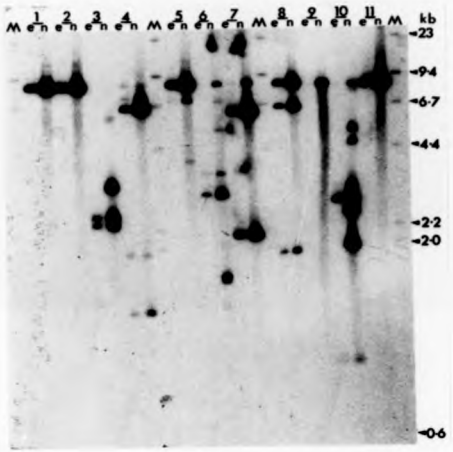

M
1 en
2 en
3 en
4 en
M
5 en
6 en
7 en
M
8 en
9 en
10 en
11 en
M
kb
23
9·4
6·7
4·4
2·2
2·0
0·6

FIGURE 12. Example of gel used to compile restriction maps of fragments containing SS and AB sequences by restriction endonuclease digestion of radiolabelled plasmid insert

The example shown is pSS47B insert. Insert DNA (100 ng) was either end-labelled with [α-^{32}P]dGTP by end-filling or labelled by nick-translation with [α-^{32}P]dGTP. Aliquots (10 ng) of end-labelled and nick-translated DNA were cleaved with various 6 bp recognition site restriction endonucleases. The restriction fragments were fractionated on a 0.8% 1x TAE agarose gel at 1.5 V/cm for 16 h. End-labelled restriction products (e) were run alongside nick-translated restriction products (n). HindIII cleaved λDNA, end labelled with [α-^{32}P]dGTP, was used as size markers (lane M): the sizes in kb denote the position of these fragments. The enzymes used are as follows:- lane 1, AvaI: lane 2, BamHI: lane 3, BglII: lane 4, EcoRI: lane 5, HindIII: lane 6, NdeI: lane7, PstI: lane 8, SalI: lane 9, SmaI: lane 9, SphI and lane 10, XhoI.

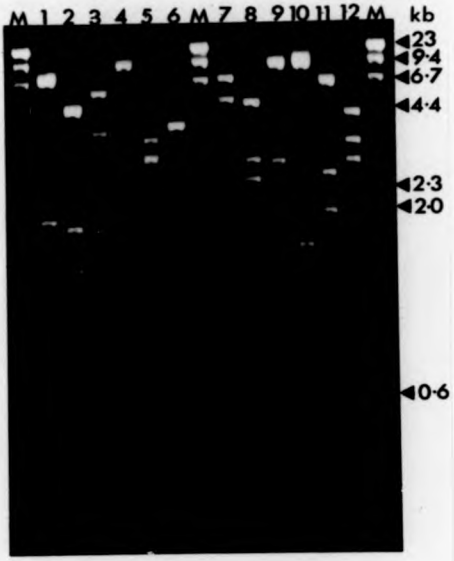

M 1 2 3 4 5 6 M 7 8 9 10 11 12 M
kb
23
9·4
6·7
4·4
2·3
2·0
0·6

FIGURE 13. Example of gel used to compile restriction maps of fragments containing SS and AB sequences by restriction endonuclease digestion of plasmid subclone DNA

Plasmid DNA (1 μg) was cleaved with various restriction enzymes and the cleavage products fractionated on a 0.8% 1x TAE agarose gel run at 1.5 V/cm for 16 h. Following electrophoresis the gel was stained with 0.1 g/ml EtBr for 15 min. HindIII cleaved λDNA was run as size markers (lane M) and the fragment sizes are shown in kb. The plasmid subclones analysed in this example, and the restriction enzyme used for cleavage are as follows:- pAB13B, PstI (lane 1); pAB13B, HindIII (lane 2); pAB13E, BamHI (lane 3); pAB13E, PstI (lane 4); pAB13E, PstI (lane 5); pSS47H, SphI (lane 6); pSS47E, SalI (lane 7); pSS47E, HindIII (lane 8); pSS47E, BamHI (lane 9); pSS47B, EcoRI (lane 10); pSS47B, SphI (lane11) and pSS47B, HindIII (lane 12).

pSS47B
pSS47E
pSS47H
BamHI
EcoRI
BamHI
HindIII
HindIII
SphI
HindIII
HindIII
SalI
BglII
SphI
HindIII
HindIII
EcoRI
PstI
BglII
EcoRI
HindIII
HindIII
SphI
HindIII
HindIII
SalI
BglII
SphI
HindIII
HindIII
EcoRI
PstI
BglII
EcoRI
HindIII
SphI
HindIII
5'
homology to
pSSU60 & pSSU160
3'
BamHI-SalI
fragment
1kb

FIGURE 14. Restriction endonuclease maps of SS47 subcloned fragments.

Restriction maps were compiled from gels similar to those shown in Figures 12 and 13. Restriction maps of the plasmid subclones pSS47B, pSS47E and pSS47H are aligned to show the overlapping regions. The region of homology to SSU60 and SSU160 is shown, as is the orientation of SS47. The scale of the map is 2 cm/kb. The BamHI-SalI fragment indicated was subjected to sequence analysis; the sequence of this fragment appears in Section III.3.B.

2.C Restriction enzyme analysis of plasmid subclones containing AB13

Restriction maps of pAB13B and pAB13E were produced as described for pSS-subclones in Section III.2.B. The same restriction enzymes were used and the maps appear in Figure 15.

2.D Discussion

The restriction fragments containing the genes of interest have been subcloned into plasmid vectors. These plasmid clones have been verified as containing SS and AB genes by restriction digestion and Southern hybridisations. The restriction maps of these clones show differences from previously characterised genes. The presence of more than one BglII insert in λSS47 does not constitute a problem since the SS gene is located at least 2 kb from any BglII site.

The SphI site characteristic of the transit-peptide/mature-polypeptide junction in all SS genes characterised to date (Bedbrook et al., 1980; Berry-Lowe et al., 1982; Broglie et al., 1983; Smith et al., 1983; Stiekema et al., 1983a; Coruzzi et al., 1984; Timko et al., 1985a; Mazur & Chui, 1985; Pichersky et al., 1986;) is present in SS47. This SphI site is separated by ~200 bp from the HindIII site in the 5'-untranslated region. This HindIII site is found in all members of the multi-gene family from pea (Cashmore, 1983; Coruzzi et al., 1984; Timko et al., 1985a), and at least one member of the soybean multi-gene family (Berry-Lowe et al., 1983). From these conserved restriction sites in the 5'-region of the gene the

pAB13B
BamH1
EcoRI
BglII
SphI
XhoI
EcoRI
PstI
BamH1
(HindIII sites not shown)
pAB13E
EcoRI
PstI
HindIII
BamH1
SphI
BglII
HindIII
EcoRI
homology to
pFab31
1kb

FIGURE 15. Restriction endonuclease maps of AB13 subcloned fragments.

Restriction maps were compiled from gels similar to those shown in Figures 12 and 13. Restriction maps of the plasmid subclones pAB13B and pAB13E are aligned to show the overlapping regions. The region of homology to pFa/b31 is shown. The scale of the map is 2 cm/kb. The HindIII sites are not shown for pAB13B; the large number of HindIII sites in this fragment prevented their accurate placement.

orientation of the gene can be concluded and is indicated in Figure 14. The presence of a 1 kb HindIII fragment containing the gene shows SS47 to differ from all previously reported pea SS genes (Cashmore, 1983; Coruzzi et al., 1984; Timko et al., 1985a). The characteristic length of this fragment, covering part of the 5'-untranslated region, the first two exons, the two introns and part of the third exon, is approximately 590 bp in the three previously reported SS genomic clones from pea (Cashmore, 1983; Coruzzi et al., 1984 and Timko et al., 1985a). The significance of the extra ~400 bp contained by this fragment in SS47 was initially unclear. It could have been due to an extra intron or a longer intron in SS47; another possibility was an extra 400 bp of coding region but this was thought to be unlikely.The position of these extra 400 bp was subsequently elucidated during DNA sequencing and shown to be in the first intron. The difference in length of the first intron between SS47 and other pea SS genes (Cashmore, 1983; Coruzzi et al., 1984; Timko et al., 1985a) may be due to intervarietal differences; SS3.6, (Cashmore, 1983), SS2.4 (Coruzzi et al., 1984) and SS8.0 (Timko et al., 1985a) were isolated from P.sativum var. Progress No.9 whereas SS47 was isolated from P.sativum var. Feltham First; this explanation is unlikely as the genes from var. Progress No.9 appear to be present in var. Feltham First (Section III.3). It is expected that SS47 has an equivalent in var. Progress no.9, possibly rbcS 3A (Fluhr et al., 1986).

From the restriction map of SS47 it was deduced that the 2.5 kb EcoRI-SalI fragment (highlighted in Figure 14), showing homology to SSU160, would contain the entire gene

and would be a suitably sized fragment for sequence analysis. The 2 kb BamHI-SalI fragment highlighted in Figure 14 shows the region of SS47 sequenced and presented in Figure 17. Further work involving the expression of SS47 in transgenic tobacco involved the use of pSS47E since this contains ~500 bp of sequence further 5' to the gene than does pSS47B. It was thought that these extra 500bp may be needed for the correct expression of this gene (Section III.6). Recent work by Nagy et al(1985). has however shown that a region of only 352 bp 5' to the transcription start site of the gene is sufficient for organ-specific and light-regulated expression.

From the restriction maps of pAB13B and pAB13E it was unclear as to the orientation of the gene within these subcloned fragments. By comparison of the maps of pAB13B and pAB13E to those of AB96 (Coruzzi et al., 1983), AB80 (Cashmore, 1983) and AB66 (Timko et al.. 1985a) it was first thought that the HindIII site indicated on the map of pAB13E (Figure 15) would be in the coding region of the gene. The orientation of the gene was resolved by sequence analysis and unexpectedly showed that the HindIII site was at the 5'-end of the gene and the BamHI site in pAB13E (the terminal BamHI site in pAB13B) was internal to the gene. The orientation of the gene also showed that the restriction map of this clone was very different to those of AB96 (Coruzzi et al.. 1983) AB80 (Cashmore, 1983) and AB66 (Timko et al., 1985a).

3. SEQUENCE ANALYSIS OF SS AND AB GENOMIC CLONES

3.A Sequencing strategy

The sequencing of SS47 and AB13 was carried out using the dideoxy-chain termination method of Sanger (1977). Either [α-^{32}P]dGTP, or [α-^{35}S]αSdATP was used (Sections II.2.K.i & ii) depending on availability.

As described in Section III.2.B, pSS47E contains the structural region of SS47 and more 5'-noncoding sequence than pSS47B. It was not known exactly how many bases upstream either clone extended so it was decided to concentrate on pSS47E. The choice of SS47E ensured that the longest available amount of 5'-noncoding sequence would be analysed. The first step in obtaining the sequence of SS47 was the production of M13 recombinant clones containing randomly generated fragments from the EcoRI-SalIsub-fragment of pSS47E highlighted in Figure 14. The random sub-fragments needed for sequencing were generated by digestion of this EcoRI-SalI fragment with various 4 bp recognition site restriction enzymes. This procedure was chosen as it would yield random overlapping fragments in the size range of a few hundred base-pairs.

The partial sequence of AB13 was obtained during the orientation of this gene within the plasmid subclones. The orientation was required to assess which plasmid clone contained the complete AB13 gene.

3.B Total sequence of SS47 genomic clone

pSS47E was first cleaved with EcoRI and the 7.5 kb EcoRI restriction fragment isolated from a 0.7% LMP agarose gel (Section II.2.E.i). The purified fragment was then cleaved with SalI and the 2.5 kb EcoRI-SalI fragment isolated from a 0.7% LMP agarose gel. This purified fragment was resuspended in SDW at 50 ng/µl. The reason for the two stage isolation process was to avoid the problem of separating the desired 2.5 kb fragment from the two 2 kb EcoRI-SalI fragments produced by cleaving pACYC184 with these two enzymes. Aliquots (5 µl) of the 2.5 kb derivative of pSS47E were digested with either MspI. Sau3A or TaqI as described in Section II.2.D.i and resuspended in SDW at 50 ng/µl. The choice of these three enzymes was based on their different 4 bp nucleotide recognition sequences and the ability to ligate their restriction products into the polylinker region of M13mp18 and M13mp19.

RF M13 vector was prepared as outlined in Section II.2.B.v. M13mp18 was cleaved with either BamHI or AccI and dephosphorylated as described in Section II.2.D.ii. The vector DNA was resuspended at 50 ng/µl in SDW. Sau3A-cleaved insert DNA was ligated into BamHI-cleaved M13mp18 DNA whilst MspI and TaqI-cleaved insert was ligated into AccI-cleaved M13mp18. Insert DNA (50 ng) was ligated into vector DNA (100 ng) as described in Section II.2.D.III. Sau3A-cleaved DNA (50 ng) was also ligated into M13mp18 DNA (100 ng) that had been cleaved with both BamHI and SalI: this ensured that the extreme 3'-region of the gene, next to the SalI site, would be represented in the fragments cloned into M13. The ligation products were transformed into E.coli strain JM103

3.B Total sequence of SS47 genomic clone

pSS47E was first cleaved with EcoRI and the 7.5 kb EcoRI restriction fragment isolated from a 0.7% LMP agarose gel (Section II.2.E.i). The purified fragment was then cleaved with SalI and the 2.5 kb EcoRI-SalI fragment isolated from a 0.7% LMP agarose gel. This purified fragment was resuspended in SDW at 50 ng/µl. The reason for the two stage isolation process was to avoid the problem of separating the desired 2.5 kb fragment from the two 2 kb EcoRI-SalI fragments produced by cleaving pACYC184 with these two enzymes. Aliquots (5 µl) of the 2.5 kb derivative of pSS47E were digested with either MspI, Sau3A or TaqI as described in Section II.2.D.i and resuspended in SDW at 50 ng/µl. The choice of these three enzymes was based on their different 4 bp nucleotide recognition sequences and the ability to ligate their restriction products into the polylinker region of M13mp18 and M13mp19.

RF M13 vector was prepared as outlined in Section II.2.B.v. M13mp18 was cleaved with either BamHI or AccI and dephosphorylated as described in Section II.2.D.ii. The vector DNA was resuspended at 50 ng/µl in SDW. Sau3A-cleaved insert DNA was ligated into BamHI-cleaved M13mp18 DNA whilst MspI and TaqI-cleaved insert was ligated into AccI-cleaved M13mp18. Insert DNA (50 ng) was ligated into vector DNA (100 ng) as described in Section II.2.D.III. Sau3A-cleaved DNA (50 ng) was also ligated into M13mp18 DNA (100 ng) that had been cleaved with both BamHI and SalI: this ensured that the extreme 3'-region of the gene, next to the SalI site, would be represented in the fragments cloned into M13. The ligation products were transformed into E.coli strain JM103

as described in Section II.2.H.i using X-gal selection for inserts into the M13 polylinker region. White plaques were picked and ssDNA prepared as outlined in Section II.2.B.vi. The number of plaques picked were as follows:- Sau3A-BamHI ligation, 50; MspI-AccI ligation, 15; TaqI-AccI ligation, 50; Sau3A-BamHI/SalI ligation, 4. These numbers were representative of the number of plaques obtained from the respective transformations. The M13mp18 recombinants were screened by T-track analysis and the following number of different clones identified: Sau3A-BamHI ligation, 12; MspI-AccI ligation, 7; TaqI-AccI ligation, 14; Sau3A-BamHI/SalI ligation, 4.

Representative clones from the various ones identified by T-track analysis were sequenced and part of the sequence of SS47 determined. The extent and orientation of the random clones used for sequence determination is indicated in Figure 16. Several clones isolated by this random method were found to contain no inserts despite producing white plaques; this is a consequence of aberrant self-ligation, or base removal by the CIP. Several other clones were found to contain sequences that did not contribute to the sequence of SS47, possibly contaminating plasmid vector fragments, or genomic fragments that align outside the region characterised. None of the MspI clones aligned with the sequence of SS47; the reasons for this are the same as for other sequences mentioned above that did not align. Regions that were not covered by these random clones were sequenced by isolating specific fragments that contained these regions; these clones are also shown on Figure 16. The entire BamHI-SalI restriction fragment

highlighted in Figure 14 was subcloned into M13mp18. Specific deletions were made by cleaving 200 ng of RF DNA from this subclone, designated mp18SS-BS, with either HindIII or SphI. The cleaved DNA was self-ligated in a volume of 50 µl. The ligation mixture was transformed into E.coli strain JM103 and white plaques picked. Single-stranded DNA was isolated from the clones picked as white plaques; these were subjected to T-track analysis.

Representative clones were sequenced. This procedure gave rise to clones extending from the BamHI site, at the 5' end of the gene, to the first and second HindIII sites further down-stream. A third clone resulting from a deletion was one containing the BamHI-SphIfragment. The clones covering the transit-peptide and first intron were isolated by purifying the 1 kb HindIII fragment from pSS47H, cleaving it with SphI and ligating the cleavage products into HindIII/SphI-cleaved M13mp19; these procedures were carried out as described above. The two possible orientations were distinguished by T-track analysis. The clones covering the third exon and 3'-untranslated region were obtained by firstly isolating the BamHI-SalI insert from mp18SS-BS. The insert was then cleaved with HindIII and the products ligated into M13mp18 cleaved with both HindIII and SalI. Out of twelve recombinants screened by T-track analysis, none contained single HindIII-SalI inserts. All contained an extra HindIII insert. The complete sequence of SS47 was obtained from analysis of all these clones (Figure 16) and covers the entire BamHI-SalI fragment highlighted in Figure 14. There was, after all, sufficient 5'-noncoding region contained in this 2 kb fragment to give the complete

5'
BamH I
Hind III
TaqI
Sau3A
HindIII
SphI
TaqI
Sau3A
Hind III
HindIII
TaqI
Sau3A
3'
Sal I
(a)
(b)
= 75 bp

FIGURE 16. Orientation of M13 subclones analysed during sequence analysis of SS47

The BamHI-SalI fragment of SS47 subjected to sequence analysis is shown. The 5'- and 3'-ends are indicated with respect to the orientation of the gene. The thick lines represent exons whilst the thin lines represent 5'-untranslated and noncoding regions, introns and 3'-untranslated and noncoding regions. Only the restriction sites used in the sequence analysis are shown. The orientation of the M13 subclones is indicated by the arrows; the length of the arrows indicates the extent of sequence obtained from the subclones. The arrow labelled (a) represents the 5'-S1 probe discussed in Section III.4.A; the arrow labelled (b) represents the 3'-S1 probe discussed in the same Section. The scale is 1 cm/75 bp.

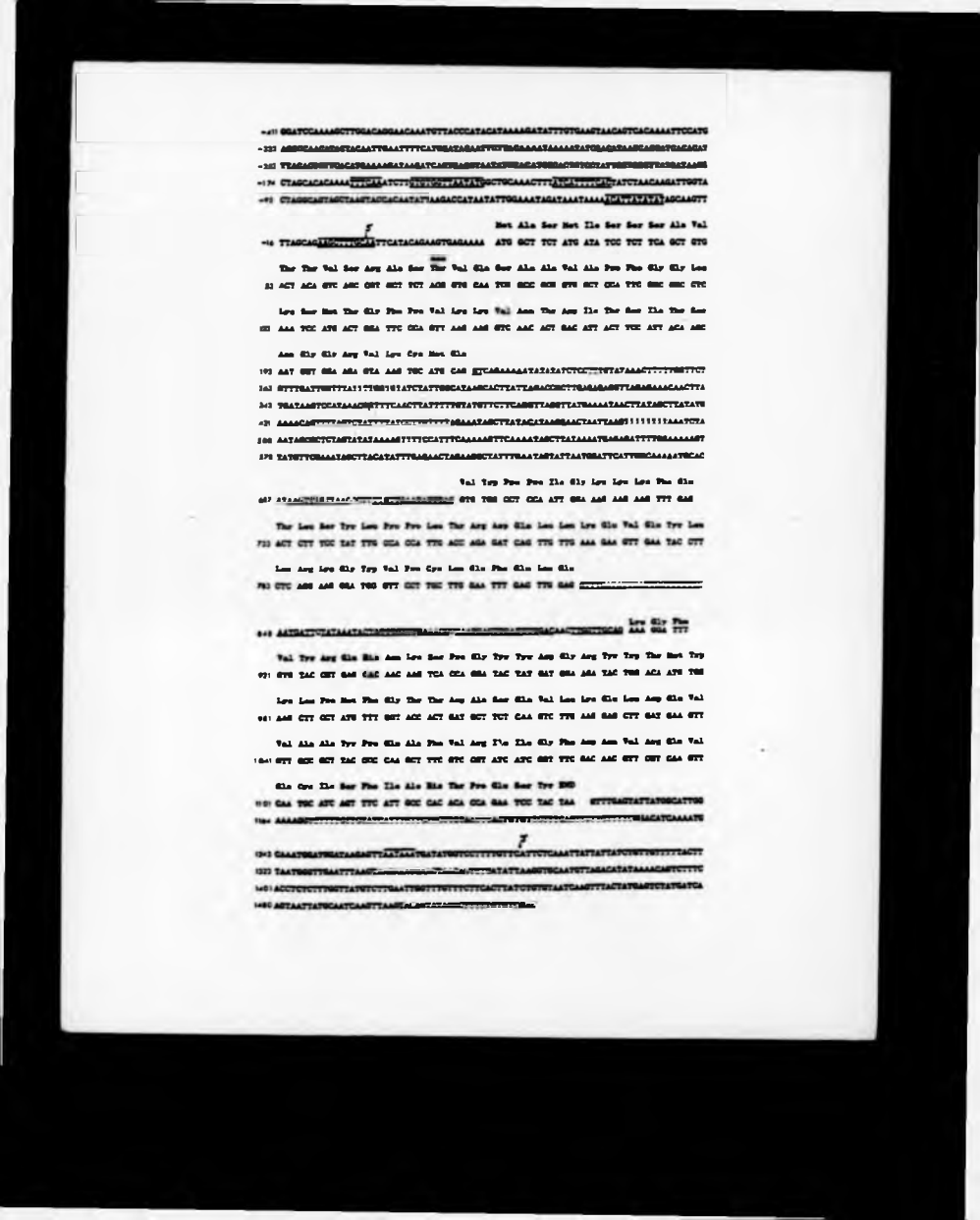


FIGURE 17. Complete nucleotide sequence of SS47

The entire 1939 bp of the BamHI-SalI restriction fragment containing the gene SS47 is presented. The sequence starts at nucleotide -411 (nucleotide number +1 represents the translation initiation point) and ends at nucleotide +1528. The strand shown corresponds to the mRNA strand. The sequence contains 433 nucleotides 5' to the initiation methionine and 383 nucleotides 3' to the stop codon. The protein sequence of this gene is presented above the exons. The translation initiation site is indicated by an arrow labelled 5' and the polyadenylation site by an arrow labelled 3'. Regions of homology to other SS genes in the 5'-untranslated region are indicated by lines above and below the sequence. Nucleotides in the introns conserved between all pea SS genes are indicated by underlining. The label END signifies the stop codon. The three asterisks (***) indicate the only unique amino acid replacement between all pea SS genes.

sequence. For this reason the remaining sequence encoded between the BamHI site and the EcoRI site at the extreme 5' end of pSS47E was not analysed. The complete sequence of the gene SS47 is shown in Figure 17.

3.C Partial sequence of AB13 genomic clone

The sequence was obtained by sub-cloning random BamHI-HindIII fragments from pAB13B into both M13mp18 and mp19. The recombinant plaques were screened by plaque hybridisation (Section II.2.G.iv) to nick-translated pFa/b31 insert DNA (Section II.2.F.i.a). The resulting positive clones, four from both M13mp18 and mp19 were screened by sequence analysis to reveal four clones in each orientation. The extent and direction of these clones is shown in Figure 18. It was postulated that if the sequence showed the 5' end of the gene to be located close to the HindIII site the BamHI site would be internal; this would in turn mean that pAB13B contained only part of the genomic sequence. If however the BamHI site was 5' to the gene both pAB13B and pAB13E would contain the entire gene. Sequence analysis revealed the former to be the case; only pAB13E contains the entire gene. The partial sequence of the AB13 gene is presented in Figure 19.

3.D Discussion

The sequence of SS47 shown in Figure 17 shows some differences to other members of the SS multi-gene family in pea as well as sharing many features of these other genes. The 5'-untranslated and 5'-noncoding region from -321 to +22 are identical to those of rbcS 3A characterised by Fluhr et

EcoRI
HindIII
BglII
SphI
BamHI
HindIII
PstI
EcoRI
=500bp

FIGURE 18. Location of AB13 within the subcloned fragment of pAB13E

The restriction map of pAB13E is shown. The thin lines represent 5'- and 3'-untranslated and noncoding regions. The thick line represents part of the AB13 exon confirmed by sequence analysis, the open continution of this line represents the predicted continuation of the exon. The 5' and 3' labels denote the orientation of the gene. The arrows below the map indicate the orientation and extent of the sequence data derived from M13 subclones. The scale is 4 cm/kb.

```
  1                                                    Met Ala Ala    Ser Ser
-29 AAGCTTATCA ATGTTTTCAA TTTCATTGCAATACA----CACC  ATG GCC GC-T ---TCA TC-------C
      A  A                                 ATA              A  CA      ATCATC
                                                          Ala Ser       SerSer

  6  Met Ala Leu Ser Phe Pro Thr Leu Thr Gly Lys Pro Val Leu Asn Pro Ser Ser
 16  ATG GCT CTC TCT TTC CCA ACC TTG ACT GGC AAG CCA GTC AAG CTG AAC CCA TCA AGC
                 CT           G          A  C
                Ser          Ala        Glu Leu

 25  Gln Glu Leu Gly Gly Ala Arg Phe Thr Met Arg Lys Ser Ala Thr Thr Lys Lys Val
 73  CAA GAA TTG GGA GGT GCA AGG TTC ACC ATG AGG AAG TCT GCT ACC ACC AAG AAA GTA
                 C
                Ala

 44  Ala Ser Ser Gly Ser Pro Trp Tyr Gly Pro Asp Arg Val Lys Tyr Leu Gly Pro Phe
130  GCC TCC TCT GGA AGC CCA TGG TAC GGA CCA GAC CGT GTT AAG TAC TTA GGC CCA TTC
   T

 63  Ser Gly Glu Ser Pro Ser Tyr Leu Thr Gly Glu Phe Pro Gly Asp Tyr Gly Trp Asp
187  TCC GGT GAG TCT CCG TCC TAC TTG ACT GGA GAG TTC CCC GGT GAC TAC GGT TGG GAC
                A

 82  Thr Ala Gly Leu Ser Ala Asp Pro Glu Thr Phe Ser Lys Asn Arg Glu Leu Glu Val
244  ACT GCC GGA CTC TCT GCT GAC CCA GAG ACA TTC TCC AAG AAC CGT GAG CTT GAA GTC

101  Ile His Ser Arg Trp Ala Met Leu Gly Ala Leu Gly Cys Val Phe Pro Glu Leu Leu
301  ATC CAC TCC AGA TGG GCT ATG TTG GGT GCT TTG GGA TGT GTC TTC CCA GAG CTT TTG

120  Ser Arg Asn Gly Val Lys Phe Ala Glu Ala Val Trp Leu Lys Ala Gly Ser
358  TCT CGC AAC GGT GTT AAA TTC GCT GAA GCT GTG TGG TTG AAG GCA GGA TCC
                          GC           C
                         Gly          Phe
```

FIGURE 19. Partial nucleotide sequence of AB13

The partial nucleotide sequence of AB13 is shown (448 bp). The sequence shown is that for the mRNA strand. The upper nucleotide strand corresponds to AB13, the lower strand to AB96 (Cashmore, 1984). The sequence of AB96 is only shown where it differs from AB13. The upper amino acid sequence is derived from AB13, the lower sequence is that derived from AB96. The AB96 sequence is only shown where it differs from that of AB13. The sequence starts at nucleotide -39 (the first nucleotide of the translation initiation site (ATG) is designated +1) and proceeds to nucleotide +409.The amino acid sequence is numbered from the initiation methionine.

al. (1986). This is the only region of the rbcS 3A gene available for comparison so it is impossible to compare the rest of SS47 with this gene. It may be that SS47 and rbcS 3A are the same gene, or they may be divergent elsewhere in their sequence. It has been noted by Mazur & Chui (1985) that the 3'-untranslated region of a SS gene from N.tabacum differs by only 2 bp in 190 from a gene from N.sylvestris yet differences, in the form of silent base substitutions, are found within the coding region. The genes AB80 (Cashmore, 1983) and AB66 (Timko et al., 1985a) share nearly 200 bp of 5'-noncoding region yet are more divergent in their coding regions and 3'-untranslated regions (Timko et al., 1985a). The inference from these observations is that the SS47 and rbcS 3A genes may differ despite a sequence homology of 100% over 343 bp; the possibility still remains however that these two genes are the same.

As noted by Fluhr et al. (1986), the 5'-untranslated region of rbcS 3A contains regions showing homology to the viral enhancer (Weiher et al., 1983) and yeast Ty element control regions (Errede et al., 1985). Due to the conserved nature of this region between rbcS 3A and SS47 these homologies are also present in SS47. The enhancer-like sequence, GTGTGGTTAA occurs at -150 in SS47 (overlined and underlined region in Figure 17). The 5'-noncoding region of SS47 also contains blocks of sequence homology, also shown as overlined and underlined regions in Figure 17, that are highly conserved between different members of the multi-gene family in pea. These regions of homology are likely to be involved in the regulation of both organ-specific and light-regulated expression (Fluhr et al., 1986). The

conserved region believed to be important for RNA polII recognition, covering the TATA box (-30), has been shown by Fluhr et al. (1986) to be involved in light-regulated expression. There is no apparent CAAT box in SS47 that shows homology to the animal consensus sequence (Benoist et al., 1980) GGC/TCAATCT or the AGGA box (C/TA$_2$-$_5$NG/TNGA$_3$-$_4$NC/TC/T) defined by Messing et al. (1983) as a possible plant equivalent of the CAAT box. The translation initiation enviroment of SS47 fits that predicted by Kozak (1981), namely ANNATGG. This sequence is thought to be involved in efficient translation initiation (Kozak, 1981).

The greatest difference between SS47 and the other pea SS genes is, as mentioned in Section III.2.D, the length of the first intron. This intron is also highly divergent in its sequence from all equivalent introns in pea SS genes. The SS genomic sequences PS2.4 (Coruzzi, 1984), SS3.6 (Cashmore, 1983) and SS8.0 (Timko et al., 1985a) all contain introns of between 75 and 83 bp for the first intron, showing very high sequence homology (~82%); the second intron is 86 bp long in all previous sequences and only 1 bp substitution is present in PS2.4 compared to SS3.6 and SS8.0. The first intron in SS47 is 471 bp long, of which only the last 31 bp show any homology to the equivalent intron in the other members of the multi-gene family. The regions of homology between introns of all members of the multi-gene family are underlined in Figure 17. The second intron in SS47 is 87 bp long and shows a high degree of homology (93%) to the equivalent introns of other pea SS genes. The intron-exon boundaries occur in the same positions in SS47 as in all other SS genes in pea (Cashmore,

1983; Coruzzi, 1984 and Timko et al., 1985a), tomato (Pichersky et al., 1986) and soybean (Berry-Lowe et al. 1982). The donor and receptor sequences for intron splicing, C/AAGGTA/GAGT and $(T/C)_nNC/TAGG$ respectively, as defined by Mount (1982), occur in the first intron in a strongly conserved form but are more divergent in the second intron. The stop codon used in SS47 is TAA and is the one used in SSU1 (Bedbrook et al., 1980), SS3.6 (Cashmore, 1983) and SS8.0 (Timko et al., 1985a).

The 3'-untranslated region of SS47, as already mentioned in Section III.1.B, shows 100% sequence homology to that of SSU60 and shows between 67% and 89% homology to the equivalent region of other members of the family. Dean et al. (1985b) have shown that different members of the SS multi-gene family in petunia can be grouped within sub-families according to homologies within the 3'-untranslated region. From the data presented in Figure 1 it is possible to assign the SS genes from pea into sub-families. The assignments are based on sequence homology and the presence of deletions within the 3'-untranslated regions. The genes PS2.4 (Coruzzi et al., 1984) and SS3.6 (Cashmore, 1983) fall within one sub-family whilst the other three genes (SSU1 [Bedbrook et al., 1980], SS8.0 [Timko et al., 1985a] and SS47 [this work]) represent three other subfamilies. It is interesting to compare the 5'-untranslated region of these SS genes and assign them into subfamilies based on this sequence data (Figure 20); if this is done a different picture emerges. SS47, SS3.6 and rbcS 3A fall within one subfamily whilst PS2.4 and SS8.0 fall within another. From this data assessed collectively,

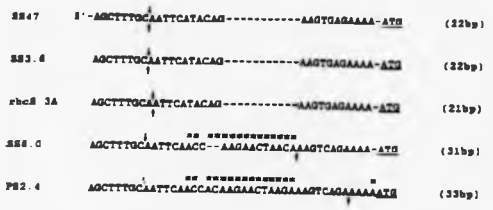

SS#7
5'-AGCTTTGCAATTCATACAG----------AAGTGAGAAAA-ATG
(22bp)
SS3.4
AGCTTTGCAATTCATACAG----------AAGTGAGAAAA-ATG
(22bp)
rbcS 3A
AGCTTTGCAATTCATACAG----------AAGTGAGAAAA-ATG
(21bp)
SS8.C
AGCTTTGCAATTCAACC--AAGAACTAACAAAGTCAGAAAA-ATG
(31bp)
PS2.4
AGCTTTGCAATTCAACCACAAGAACTAAGAAAGTCAGAAAAAATG
(33bp)

FIGURE 20. Sequence comparison between the 5'-untranslate regions of SS47 and other pea SS genes

The 5'-untranslated region of SS47 is shown with other SS genes, namely, SS3.6 (Cashmore, 1983), rbcS 3A (Fluhr and Chua, 1986), SS8.0 (Timko et al., 1985a) and PS2.4 (Coruzzi et al., 1984), aligned between the conserved HindIII site at the 5'-end and the translation initiation codon. A dash (-) denotes a missing base and an asterisk (*) denotes a base change as compared to SS47. The initiation methionine is underlined. The figures in brackets indicate the length of the 5'-untranslated region. Downward pointing vertical arrows indicate the transcription start site and upward pointing arrows indicate the S1 cleavage site (based on a 2 bp mismatch) in a RNA-DNA hybrid between the SS47 5' region and RNA from other members of the family.

it appears that PS2.4 has a 5'-untranslated region region from one subfamily and a 3'-untranslated region from another; this is a possible consequence of either gene duplication and subsequent recombination, gene conversion, or less room for divergence in the 5'-untranslated region than in the 3'-untranslated region. The polyadenylation signal in SS47 conforms to the consensus AATAAA. It is the only pea SS gene so far analysed to contain the sequence in its consensus form. This sequence is however not unknown in plants (Messing et al., 1983) but appears to be functional in various slightly modified forms (Messing et al., 1983). The site of polyadenylation in Figure 17 is indicated by a vertical arrow labelled 3'. This site occurs 19 bp 3' to the end of the polyadenylation signal.

The coding region of SS47 shows a very high degree of sequence homology to the other members of the family. A few silent base-substitutions occur in the coding region and some base changes occur in the transit-peptide that alter the amino acid sequence of this peptide. Threonine^{18} is highlighted in Figure 17 as the only unique amino acid replacement compared to the other genomic sequences analysed. The threonine replaces an arginine. In SS47 valine^{18} replaces a glycine found in PS2.4 (Coruzzi et al., 1984) and SS3.6 (Cashmore, 1983); this conservative substitution is also present in SS8.0. In a comparison of SS47 to SSU1 there are 3 amino acid changes in the last 13 amino acids of the transit-peptide (the entire transit-peptide sequence of SSU1 is not available) and 9 amino acid substitutions in the coding region. This variation is surprising due to the highly conserved nature

of the amino acid sequence within other members of the family (Cashmore, 1983; Coruzzi et al., 1984 and Timko et al., 1985a). The lack of variability within the coding region of genes from different varieties, namely SS47 (var. Feltham First) and PS2.4 (Coruzzi et al., 1984), SS3.6 (Cashmore, 1983) and SS8.0 (Timko et al., 1985a) (var. Progress No.9) makes the differences between SSU1 and SS47 even more surprising since they were both isolated from the same variety of pea. It has been noted by Mazur and Chui (1985) that the transit-peptides of SS genes show far greater variability in their amino acid composition than do the mature polypeptides; this observation holds for pea with the exception of SSU1. There are some third position changes in some codons in the mature sequence of SS47 as compared to PS2.4 (Coruzzi, 1984), SS3.6 (Cashmore, 1983) and SS8.0 (Timko et al., 1985a) as there are between these three sequences themselves. These changes do not alter the protein sequence as has been discussed above.

The 16 amino acid block highlighted by Broglie et al. (1983); Mazur and Chui (1985) and Shinozaki and Sugiura (1983) as being conserved between different species is found in SS47 between tyrosine^{118} and glycine^{133}. This sequence, namely:- Tyr Tyr Asp Gly Arg Tyr Trp Thr Met Trp Lys Leu Pro Met Phe Gly, is found in all higher plant SS genes analysed to date (Bedbrook et al., 1980; Berry-Lowe et al., 1982; Broglie et al., 1983; Cashmore, 1983; Dunsmuir et al., 1983a; Stiekema et al., 1983a; Pinck et al., 1983; Coruzzi et al., 1984; Timko et al., 1985a; Mazur and Chui, 1985; Pichersky et al., 1986). The high degree of conservation of this region suggests a functional role in the binding of the

highly conserved LS.

In summary, the sequence presented in Figure 17 is different in several respects from any previously described from pea; it may however be the entire sequence of which Fluhr et al. (1986) presented the extreme 5'-end. Despite showing several differences to other pea SS genes, it shows a high level of conformity to the structural pattern emerging for the SS multi-gene family from pea.

The partial AB sequence presented in Figure 19 is the 5'-end of AB13. This sequence extends from nucleotide -39, numbered from the translation start site, to nucleotide 408. The most striking differences between AB13 and the other AB genes from pea, AB96 (Broglie et al. 1983), AB80 (Cashmore, 1983) and AB66 (Timko et al., 1985a), are the presence of a HindIII site in the 5'-untranslated region and a BamHI site in the coding region. Both the HindIII site and the BamHI site in AB13 are produced by single base substitutions from AB96. The sustitution that produces the BamHI site does not affect the amino acid encoded by the relevant codon. The 5'-untranslated region that has been sequenced does not extend far enough upstream to encompass the 5'-regulatory elements discussed in relation to SS47. This region does however show a high degree of homology (87%) to the corresponding regions of AB80 and AB66 (Timko et al., 1985a). The environment around the first methionine is a high efficiency translation-initiation environment (Kozak, 1981). There are 9 bp deleted in AB13 compared to the other three pea AB genes characterised. These deleted bases do not interupt the reading frame and have the sole effect of deleting three of the five serine residues present

in the transit-peptide of both AB80 and AB66 (Timko et al., 1985a). There are 5 amino acid changes due to base substitutions in the transit-peptide and two in the mature polypeptide. The greater heterogeneity shown between transit-peptides of various members of the family than between mature polypeptides is not surprising since this variation has been observed in the AB genes of petunia (Dunsmuir, 1985) and has been discussed in relation to the SS gene families above. The homology at the nucleotide level of AB13 to AB80 and AB66 is closer than that shown for AB80 to AB96 (Cashmore, 1984). It is however not as closely related to either AB80 or AB66 as they are to each other (Timko et al., 1985a). Both AB80 and AB66 share identical amino acid sequences in their transit and mature polypeptides.

The differences observed between AB13 and the other members of the family are not surprising since similar variation has been observed in the AB genes of Arabidopsis thaliana (Leutwiler et al., 1985) where one AB gene, AB140, differs from another, AB165, by one amino acid in the transit-peptide; in petunia (Dunsmuir, 1985) different members of the multi-gene family show differences at both the nucleotide and polypeptide levels. In agreement with all AB genes analysed to date, with the exception of AB19 from Lemna (Karlin-Neumann et al., 1985), AB13 contains no introns in the region sequenced. In conclusion it can be stated that AB13 encodes an AB sequence that differs from other pea AB genes characterised. It should however be noted that the differences between AB13 and the other pea AB genes may in part be due to intervarietal differences (Section

III.4.D). AB13 was isolated from P.sativum var. Feltham First whilst the other three AB genes were isolated from P.sativum var. Progress no.9, although intervarietal differences do not seem to be great within the SS gene family from these two varieties. It is possible that AB13 is an expressed member of the gene family due to the homology in the part of the 5'-untranslated region analysed to other AB genes. The presence of an open reading frame coding for a polypeptide of conserved sequence also indicates the possible activity of this gene. In Section III.4 S1 nuclease mapping data is presented that indicates the transcriptional activity of AB13.

4. EXPRESSION OF SS47 AND AB13 IN PISUM SATIVUM

4.A Introduction

To identify the transcriptional start site of SS47, an [α-^{32}P]dGTP-labelled genomic DNA fragment extending beyond the expected transcriptional start site of the gene was annealed to total pea RNA. S1 nuclease digestion of RNA-DNA hydrids was performed and the protected fragments sized on polyacrylamide gels (Section II.2.C.iii.a). The probe used is shown in Figure 21. This probe is an M13 sub-clone of SS47 and was isolated during sequence analysis of SS47; it consists of a 204 bp HindIII-SphI fragment cloned into M13mp18. The fragment is from the 5'-end of the gene and contains 174 nucleotides of coding sequence and 30 nucleotides 5' to the initiator methionine. By comparison to the 5'-ends of other pea SS genes (Figure 20) it was concluded that 30 nucleotides 5' to the coding region would be sufficient to encompass the transcriptional start site. The anti-sense DNA strand was synthesised from the M13 ssDNA clone as outlined in Section II.2.F.iii. The probe was such that on annealing to homologous mRNA species, 79 nucleotides at the 5'-end of the probe (composed entirely of M13 sequence), would remain single-stranded and would therefore be removed by the S1 nuclease. The number of nucleotides removed from the 3'-end of the probe would depend on its homology to the various members of the SS gene family represented in the mRNA. A 3 bp consecutive mismatch is sufficient for the S1 nuclease to create a nick in the RNA-DNA hybrid (Dodgson and Wells, 1977). Despite the amino

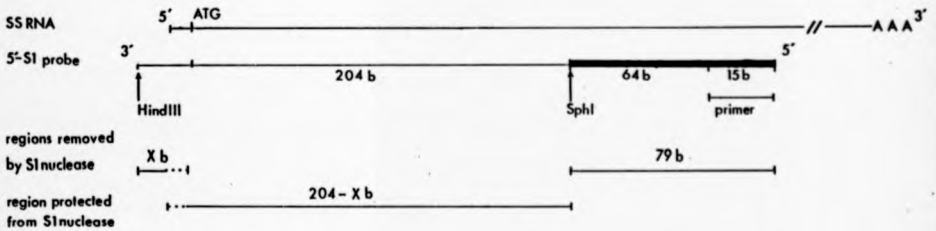

SS RNA
5′
ATG
AAA 3′
5′-S1 probe
3′
5′
204 b
64 b
15 b
HindIII
SphI
primer
regions removed
by S1 nuclease
X b
79 b
region protected
from S1 nuclease
204 – X b

FIGURE 21. SS47 5'-S1 probe

The S1 probe covers 204 b of the 5'-end of SS47 extending from the HindIII site 5' to the transcription initiation site to the SphI site at the transit peptide/mature polypeptide junction (thin line) (Figure 16). The fragment is cloned into M13mp18 (thick line). The anti-sense probe is produced by extension of the M13 sequencing primer annealed to ssDNA. Cleavage with HindIII generates an anti-sense strand 283 b long of which 79 b are M13 sequence. The 3'- and 5'- ends of the probe are indicated. The SS RNA is shown aligned to the probe. The relative position of the initiation ATG is indicated as are the 5'- and 3'-ends of the message. The regions removed from an RNA-DNA hybrid are shown; 79 b from the 5'-end of the probe and X b from the 3'-end of the probe. The value of X is dependent on the homology between the 5'-untranslated region of SS47 and other SS RNA molecules present. The predicted cleavage points for the other SS genes are shown in Figure 20. The region protected from S1 nuclease digestion is shown. The predicted size of this fragment for the different members of the SS multi-gene family is shown in Table 4.

acid substitutions in the transit-peptide of SS47 as compared to the other members of the family (Section III.3.D), no more than one consecutive base mismatch occurs within the 174 nucleotides of coding region of the probe and the corresponding region from the other members of the family. Assuming no intervarietal differences, the protected region of the probe for all members of the family thus far characterised will therefore be at least 174 nucleotides. The assumption of no intervarietal differences is a valid one as will be seen later (this Section).

The 5'-untranslated regions of SS47, SS3.6 (Cashmore et al., 1983) and rbcS 3A (Fluhr et al., 1986) are identical to one another, as are the corresponding regions of PS2.4 (Coruzzi et al., 1984) and SS8.0 (Timko et al., 1985a). The expected S1 cleavage points for these SS genes are shown in Figure 20. From this comparison of sequences, the protected fragments are expected to be of at least three size classes, corresponding to mRNA from SS47, SS3.6 and rbcS 3A and from SS8.0 and from PS2.4. Other sizes of protected fragments will correspond to as yet uncharacterised members of the SS multi-gene family. The sizes of expected 5'-S1 protected fragments are shown in Table 4. The S1 analysis procedure was carried out as described in Section II.2.J.i using total RNA from 9-day-old leaves of P.sativum var. Feltham First. The use of a uniformly labelled single-stranded probe was chosen in preference to an end-labelled probe because of its convenience and high sensitivity (Burke, 1984).

The 3'-end of SS47 (arrowed in Figure 17) was assigned by comparison to the poly(A)$^+$ tail addition point of SSU60. S1 nuclease analysis of the 3'-end of SS genes was

TABLE 4. Predicted fragment sizes from 5'-S1 nuclease analysis of the pea SS multi-gene family

The predicted protected fragment sizes are shown for the different members of the pea SS multi-gene family following S1 nuclease analysis with the probe shown in Figure 21.

	nucleotides
SS47	197
SS3.6	197
rbcA 3A	196
SS8.0	186
PS2.4	179

TABLE 5. Predicted fragment sizes from 3'-S1 nuclease analysis of the pea SS multi-gene family

The predicted protected fragment sizes are shown for the different members of the pea SS multi-gene family following S1 nuclease analysis with the probe shown in Figure 22.

	nucleotides
SS47	230
SS3.6	159
PS2.4	159
SS8.0	159
SSU1	88

performed to examine the steady state levels of expression of SS47 in relation to other members of the family. The 5'-S1 nuclease analysis of SS genes will not distinguish between SS47 and several other members of the family, whereas a comparison of the 3'-untranslated regions of the SS genes (Figure 1) indicates that 3'-S1 analysis will. The probe used for 3'-S1 analysis is shown in Figure 22. The probe is produced from a 797 bp Sau3A fragment from SS47 cloned into M13mp19. This fragment extends from 332 bp 3' to the stop codon, through the 3rd intron, 2nd exon and half of the 2nd exon (see Figure 16 label b). The orientation of the fragment in M13 ensures that the single-stranded DNA produced by extension of the M13 sequencing primer is complementary to the mRNA. During the preparation of the probe, (Section II.2.J.i) the in vitro-synthesized double stranded DNA was cleaved with HindIII, denatured and the single-stranded labelled fragment isolated from a gel. This HindIII cleavage yields a product of 447 nucleotides covering 332 nucleotides of 3'-untranslated and non-coding region and 86 nucleotides of coding region extending to the last HindIII site in the gene (see Figure 22). When this probe is annealed to SS mRNA, and subjected to S1 nuclease digestion, 246 nucleotides from the 5'-end of the probe will be removed (Figure 22) These 246 nucleotides contain M13 sequence and 3'-non-coding sequence. A fragment of 231 nucleotides will be protected by SS47 mRNA. Other genes will, depending on their homology to the 3'-untranslated region of SS47, produce shorter protected fragments. From comparisons of the 3'-untranslated regions of all known pea SS genes (Figure 1) the S1 cleavage points can be predicted.

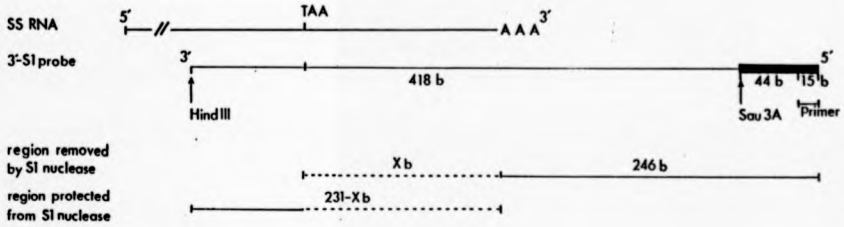

SS RNA
5'
TAA
A A A
3'
3'-S1 probe
3'
418 b
5'
44 b
15 b
Hind III
Sau 3A
Primer
region removed by S1 nuclease
X b
246 b
region protected from S1 nuclease
231-X b

FIGURE 22. SS47 3'-S1 probe

The S1 probe covers 418 b of the 3'-end of SS47 extending from the last HindIII site in the coding region (Figure 16) to the last Sau3A site in the BamHI-SalI fragment (thin line). The fragment is cloned into M13mp18 (thick line). The anti-sense probe is produced by extension of the M13 sequencing primer annealed to ssDNA. Cleavage with HindIII generates an anti-sense strand 477 b long of which 59 b are M13 sequence. The 3'- and 5'-ends of the probe are indicated. The SS RNA is shown aligned to the probe. The relative position of the translation termination codon (TAA) is indicated as are the 5'- and 3'-ends of the message. The regions removed from an RNA-DNA hybrid are shown; 246 b from the 5'-end of the probe (59 b of M13 sequence and 187 b of 3'-noncoding region) and X b from the 3'-end of the probe. The value of X is dependant on the homology between the 3'-untranslated region of SS47 and other SS RNA molecules present. The predidicted cleavage points for the other SS genes are shown in Figure 1. The region protected from S1 nuclease digestion is shown. The predicted sizes of this fragment for the different members of the SS multi-gene family are shown in Table 5.

Again this argument assumes that no intervarietal differences occur. The predicted sizes of the 3'-S1 protected fragments are shown in Table 5.

Steady state levels of expression of different members of the SS multi-gene family in pea were examined in different organs of the plant as well as in dark-grown leaf tissue. The 3'- and 5'-S1 probes discussed above were used.

The expression of AB13 was examined in light and dark grown 9-day-old leaf tissue as well as in root and stem tissue. S1 nuclease protection experiments, as used for SS above, were carried out using the probe shown in Figure 23. This probe is a 447 bp HindIII-BamHi fragment encoding the 5'-end of AB13 (Figure 18). The fragment was cloned into M13mp19: in this orientation the product extended from the sequencing primer is complementary to the mRNA. This fragment does not extend beyond the transcriptional start site (as predicted by comparison to AB80 and AB66 [Timko et al., 1985a]). The probe contains 59 nucleotides of M13 sequence which will remain as single-stranded DNA once the probe is annealed to AB mRNA. These 59 nucleotides will be removed by S1 nuclease to yield a protected fragment of 447 nucleotides corresponding to AB13 mRNA: other members of the family that show divergence in their 5'-untranslated regions will yield protected fragments of less than 447 nucleotides. It was expected that, as for the SS multi-gene family, the AB multi-gene family will produce different-sized protected fragments corresponding to the different members of the family. By comparison of the 5'-end of AB13 to AB80 and AB66 (Timko et al., 1985a) it was predicted that AB80 and AB66 would produce protected fragments the same size as each other (399 nucleotides) and 48 nucleotides less than the

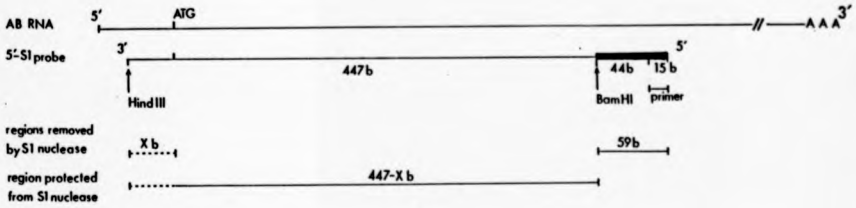

AB RNA
5'
ATG
//
A A A
3'
5'-S1 probe
3'
5'
447b
44b
15b
Hind III
BamHI
primer
regions removed by S1 nuclease
X b
59b
region protected from S1 nuclease
447-X b

FIGURE 23. AB13 5'-S1 probe

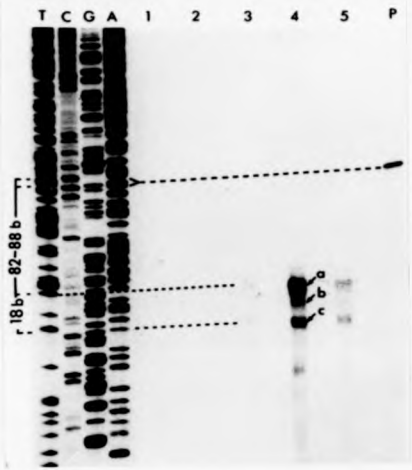
The S1 probe covers 447 b from the 5'-end of AB13 extending from the HindIII site 5' to the first methionine to the BamHI site internal to the gene (thin line) (Figure 16). The fragment is cloned into M13mp18 (thick line). The anti-sense probe is produced by extension of the M13 sequencing primer annealed to ssDNA. Cleavage with HindIII generates an anti-sense strand 506 b long of which 59 b are M13 sequence. The 3'- and 5'-ends of the probe are indicated. The AB RNA is shown aligned to the probe. The relative position of the initiation ATG is indicated as are the 5'- and 3'-ends of the message. The regions removed from an RNA-DNA hybrid are shown; 59 b from the 5'-end of the probe and X b from the 3'-end of the probe. The value of X is dependent on the homology between the 5'-untranslated region of AB13 and other AB RNA molecules present. The region protected from S1 nuclease digestion is shown.

fragment corresponding to AB13. These fragment sizes were deduced from the comparison of AB13 and AB80 (Figure 19).

The probes used in the S1 analysis of both the AB and SS genes showed no homology within the M13 region of the probe to either AB or SS sequences. This was particularly important over the M13/insert junction where M13 nucleotides might correspond to nucleotides within the coding region adjacent to the probe, thereby protecting extra bases from S1 nuclease digestion.

4.B Transcriptional start and stop sites of SS47

The conditions for annealing the single-stranded DNA probe (shown in Figure 21) to RNA for S1 analysis were established as follows. Pea leaf total RNA (1 µg) from 9-day-old light-grown plants (Section II.2.B.ix) was mixed with 20 pg single-stranded prime-cut probe (Section II.2.F.iii). Replicate RNA-probe annealing mixes were annealed for 3 h at temperatures between 55°C and 70°C at 5°C intervals. Pea total RNA (10 µg) was also annealed to 20 pg prime-cut probe at 65°C for 3 h. The RNA-DNA hybrids were digested with S1 nuclease as described in Section II.2.J.i and the nuclease-protected fragments analysed on a 6% denaturing polyacrylamide gel (Section II.2.C.iii.a). The gel was exposed to X-ray film overnight; this gel appears in Figure 24. From these results, the optimal annealing temperature was ascertained as 70°C and the transcriptional start site of SS47 tentatively assigned. As mentioned in the above section, the transcription termination site of SS47 was identified from the site of polyadenylation in the corresponding cDNA, namely SSU160.


T C G A 1 2 3 4 5 P
r 18b — 82–88 b
a
b
c

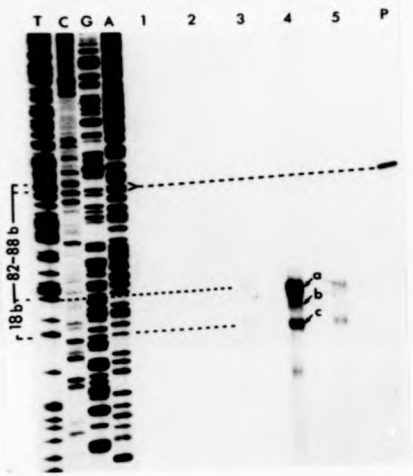

T C G A 1 2 3 4 5 P
82~88 b
18b
a
b
c

FIGURE 24. 5'-S1 analysis of SS genes in pea

The conditions for RNA-DNA annealing were characterized for the 5'-S1 probe (Figure 21). Total leaf RNA from 9-day old light-grown pea plants (1 μg) was annealed to 20 pg 5'-S1 probe at 55°C (lane 1), 60°C (lane 2), 65°C (lane 3) and 70°C (lane 5). RNA (10 μg) was annealed to 20 pg 5'-S1 probe at 65°C (lane 4). The annealed RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on 6% denaturing polyacrylamide gels at 40 W for 1.5 h and the gels exposed to X-ray film overnight. The protected fragments labelled a, b and c correspond to different members of the multi-gene family (see text). The S1 probe (1.5 pg) is shown in the lane labelled P. M13 sequence reactions of the S1 probe were run as size markers. The reactions are labelled according to the ddNTP present. The size difference between the probe and fragment a is shown, as is the difference in size between fragment a and fragment c. The number of bases was determined from a low exposure autoradiograph of the gel and by proof-reading from previous sequencing of this fragment.

4.C Organ-specific expression of SS47

The expression of SS genes in various organs was investigated, as well as the levels of expression in dark-grown leaf tissue, by S1 analysis. The S1-probe described in Section III.4.A (Figure 21) was used for these investigations. The probe was labelled as described in Section II.2.F.iii. Total pea leaf RNA (1, 5 or 10 μg) was annealed to 50 pg prime-cut probe (section II.2.F.iii) at 70°C for 3 h. The use of increasing amounts of RNA was to ensure that conditions of probe-excess were in use. Conditions of probe-excess were necessary to ensure quantitative results. Total RNA (10 μg) from 9-day-old light-grown stem tissue, root tissue and 9-day-old dark-grown pea apices were each annealed with 50 pg of the above probe. Xenopus borealis tadpole RNA (10 μg) (a gift from C.Wilson) was also annealed to 50 pg of probe as a control to identify any nuclease-protected fragments that may be due to probe secondary structure or homology to non-SS sequences; none were observed (Figure 25, lane 7). Following the 3 h annealing, the RNA-DNA hybrids were digested with S1 nuclease (Section II.2.J.i) and the nuclease-protected fragments fractionated on 6% polyacrylamide gels (Section II.2.C.iii.a). End-labelled HpaII digestion products of pBR322 (a gift from A.Mohamed) were used as size markers on the gel. The levels of products from 1, 5 and 10 μg of pea leaf RNA indicated that the S1 analyses were quantitative. The gel is shown in Figure 25. No expression was detected in stem or root tissue (Figure 25, lanes 4 and 5). Low levels of expression were observed

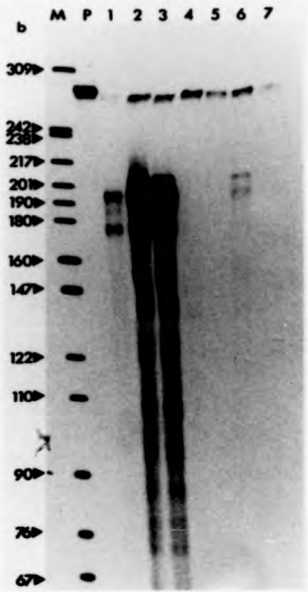

b
M
P
1
2
3
4
5
6
7
309
242
238
217
201
190
180
160
147
122
110
90
76
67

FIGURE 25. Organ specific expression of SS genes in pea

Total RNA (1, 5 and 10 μg) from 9-day-old pea plants grown in the light was annealed to 50 pg of single-stranded 5'-specific S1 probe as described in Figure 21. Total RNA (10 μg) from 9-day old, light-grown roots and stems was each annealed to 50 pg single-stranded probe as was 10 μg total RNA from 9-day old dark grown pea apices. Annealing was performed at 70°C for 3 h. The RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on a 6% denaturing polyacrylamide gel at 40 W for 1.5 h. The gel was exposed to X-ray film overnight. The size markers in lane M are end-labelled HpaII digestion products of pBR322 (a gift from A.Mohamed); the sizes are shown in bases. The probe (10 pg) was run in lane P. Lanes 1, 2 and 3 represent 1, 10 and 5 μg total RNA respectively. Stem RNA, root RNA and dark-grown RNA products were run in lanes 4, 5 and 6 respectively. Lane 7 contains 10 μg Xenopus borealis tadpole RNA annealed to 50 pg S1 probe and digested with S1 nuclease as a control. The over-exposure of lanes 2 and 3 was necessary to visualise all the bands in lane 6; lower exposure of lanes 2 and 3 reveal the same banding pattern found in lane 1.

in dark-grown pea apices (Figure 25, lane 6).

4.D Organ-specific expression of AB13

To investigate the level of expression of AB13 in different organs of pea, total RNA from different organs was annealed to a single-stranded DNA S1 probe (Figure 23) and subjected to S1 nuclease digestion. Total RNA (1, 5 or 10 μg) from 9-day-old light-grown leaf tissue was annealed to 50 pg of prime-cut probe (Section II 2.J.i) at 70°C for 3 h. RNA (10 μg) from 9-day-old dark-grown pea apices, RNA (10 μg) from 9-day-old light-grown stem tissue and RNA (10 μg) from 9-day-old root tissue from light-grown plants was also annealed to prime-cut probe as described above. X.borealis tadpole RNA (10 μg) was also annealed to 50 pg of probe as a control (Section III.4.D). The RNA-DNA hybrids were digested with S1 nuclease as described in Section II.2.J.i and the protected fragments analysed on a 6% denaturing polyacrylamide gel (Section II.2.C.111.a); the gel is shown in Figure 26. Pea leaf RNA (1, 5 and 10 μg) in lanes 1, 2 and 3 of Figure 26 indicate a quantitative increase in the signal as the RNA level is increased: the signals seen in other lanes can therefore be correlated to those seen in the first three lanes. Only one S1 product is seen; the S1 probe used does therefore not distinguish between different members of the multi-gene family. Dark grown-leaf tissue shows the presence of AB mRNA at a level approximately equal to one-tenth of that found in light grown leaf-tissue whereas stem and root tissue show no detectable AB mRNA. From the X.borealis tadpole RNA control (Figure 26, lane 9) it can be seen that none of the bands seen are due to either

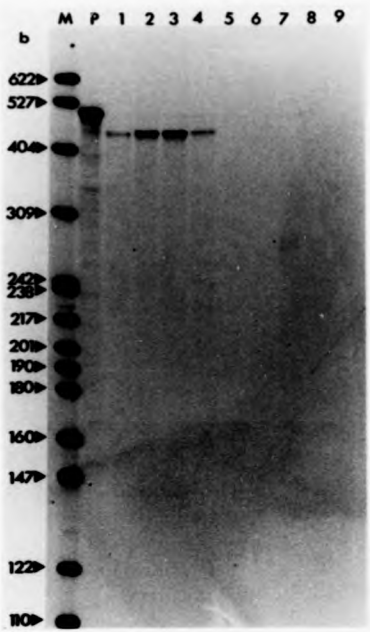

M P 1 2 3 4 5 6 7 8 9
b
622
527
404
309
242
238
217
201
190
180
160
147
122
110

FIGURE 26. Organ specific expression of AB genes in pea

Total RNA (1, 5 and 10 μg) from 9-day-old pea plants grown in the light was annealed to 50 pg of single-stranded 5'-specific S1 probe as described in Figure 23. Total RNA (10 μg) from 9-day old, light-grown roots and stems was each annealed to 50 pg single stranded probe as was 10 μg total RNA from 9-day old dark grown pea apices. Annealing was performed at 70°C for 3 h. The RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on a 6% denaturing polyacrylamide gel at 40 W for 1.5 h. The gel was exposed to X-ray film overnight. The size markers in lane M are end-labelled HpaII digestion products of pBR322 (a gift from A.Mohamed); the sizes are shown in bases. The probe (10 pg) was run in lane P. Lanes 1, 2 and 3 represent 1, 5 and 10 μg total RNA respectively. Dark-grown RNA, stem RNA and root RNA S1 products were run in lanes 4, 5 and 6 respectively. Lane 9 contains 10 μg Xenopus borealis tadpole RNA annealed to 50 pg S1 probe and digested with S1 nuclease as a control. The samples loaded in lanes 7 and 8 are discussed in Figure 41.

secondary structure or hybridisation to non-AB sequences.

4.D Discussion

The annealing conditions for RNA-probe annealing were chosen from the data presented in Figure 24. Annealing temperatures of 55°C and 60°C appear to give very low levels of annealing over 3 h and subsequent low levels of protection from S1 nuclease. (Figure 24, lanes 1 and 2). At 65°C, 1 µg of RNA (Lane 3) gives faint bands corresponding to protected fragments of ~197 b (a) and one of ~179 b (c). These bands correspond to probe molecules that have lost, in addition to the 79 b of M13 sequence from their 5'-end, ~8 b and ~26 b respectively from their 3'-ends. In Lane 4 (10 µg RNA annealed at 65°C) these bands can be seen in more clearly; a third band (b) corresponding to a protected fragment of ~186 b is also visible. This band corresponds to probe molecules that have lost ~19 b from their 3'-end. The optimum temperature for annealing is however 70°C (lane 5) as can be seen from the highest levels of protected fragments from 1 µg RNA.

As discussed in Section III.3.A, sequence comparisons have been made between four SS genes isolated from P.sativum var. Progress No.9 (PS2.4 [Coruzzi et al., 1984], SS3.6 [Cashmore, 1983], SS8.0 [Timko et al., 1985a] and rbcS 3A [Fluhr et al., 1986]) and SS47 within the 5'- and 3'-untranslated regions. The sizes of 5'-fragments protected from S1 nuclease were predicted for the above genes in Table 4. The predicted size for the protected fragment corresponding to SS47 was 197 b. Band a in Figure

24 (~197 b) fits well with this prediction. From the intervarietal comparisons it was also predicted, if equivalent genes are found in Feltham First, that SS3.6 and rbcS 3A would also yield products of 197 b. SS8.0 was expected to yield a protected band of 186 b. A faint ~186 b band is visible in Figure 24, lane 4. PS2.4 was predicted to produce a protected fragment of 179 b; this corresponds to band c in Figure 24. The sizes of the protected fragments were calculated from the sequence markers on the gel shown and by comparison of this sequence to other sequencing reactions of the same fragment. Figure 25 shows the same S1 protected fragments fractionated alongside [^{32}P]-labelled pBR322 HpaII digestion products. The presence of bands of the predicted sizes for the genes from Progress No.9 indicate the presence of homologous genes in Feltham First. The assumption that intervarietal size comparisons can be made is therefore a correct one. Conclusions can thus be drawn about the levels of expression of various members of the multi-gene family; these will be discussed later in this section.

From the data presented in Figures 24 and 25, it can be concluded that the 5'-cap site is located between the nucleotide designated +1 (indicated by an arrow in Figure 17) and the nucleotide designated +6. Due to the spread of the S1 product in Figure 24, lane 4, the corresponding sequence band cannot be assigned beyond all doubt. However, by comparison to other SS genes, the base designated +1 corresponds to the cap site found in PS2.4, SS3.6, and SS8.0, whilst +2 in SS47 corresponds to the designated cap site of rbcS 3A. The cap site always occurs at an adenine

(Breathnach and Chambon, 1981); the only possible location of it within the +1 to +6 region is at either +1, +2 or +6. From the data obtained during this work it is not possible to define the cap site in any greater detail; the numbering of SS47 is however consistent with the cap site located at +1 (the site used in the majority of other SS genes and the probable site in SS47).

The assignment of the polyadenylation site was made from the sequence of SSU60, the corresponding cDNA to SS47. The S1 protected fragments obtained from 3'-S1 analysis of the SS genes (Figure 27) could not be used to assign the polyadenylation site since sequence markers were not run for this gel. The doublet at around 230 b corresponds to the length of the protected fragment predicted for SS47. The doublet is due entirely to SS47 mRNA and not to two different members of the family; this is concluded from the presence of the same doublet in the S1 analysis of RNA isolated from transgenic tobacco expressing SS47 and containing no other pea genes (Figure 42). The doublet, rather than a single band, may be due to artefactual S1 digestion; the lower band being the product of anomylous digestion of the RNA-DNA duplex, due to breathing of the duplex, at an A/T rich region. An 8 b T-rich region is found 5' to the predicted polyadenylation site (arrowed and labelled 3' in Figure 17). The difference in size of the two members of the doublet is ~8 bp, the size difference that would be caused by the breathing of the T rich region (see Figure 17). This artifact would however be expected to give a ladder effect between the two bands of the doublet; this ladder is not seen. The other possible explanation for the

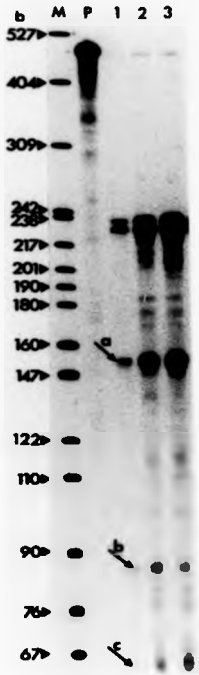

b
M
P
1
2
3
527
404
309
242
238
217
201
190
180
160
147
122
110
90
76
67
a
b
c

FIGURE 27. 3'-S1 analysis of SS genes in pea

Total RNA (1, 5 and 10 µg) from 9-day-old pea plants grown in the light was annealed to 50 pg of single-stranded 3'-specific S1 probe as described in Figure 22. Annealing was performed at 70°C for 3 h. The RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on a 6% denaturing polyacrylamide gel at 40 W for 1.5 h. The gel was exposed to X-ray film overnight. The size markers in lane M are end-labelled HpaII digestion products of pBR322 (a gift from A.Mohamed); the sizes are shown in bases. The probe (10 pg) was run in lane P. Lanes 1, 2 and 3 represent 1, 5 and 10 µg RNA respectivly. The doublet at around 235 b corresponds to SS47 mRNA, The bands labelled a and b correspond to other members of the SS multi-gene family. The band labelled c is an S1 artefact band (see text). The faint bands between the main bands are due to either weakly expressed SS genes or breakdown products of the S1 probe.

doublet is a second polyadenylation site. Such a second site would, from the intensities of the bands, be expected to be used approximately 50% of the time.

The investigation into the expression of different members of the SS multi-gene family in different organs revealed that no detectable mRNA was present in either stem or root tissue (Figure 25). The S1 technique used with uniformly-labelled DNA probes is 100x as sensitive as S1 analysis with end-labelled probes (Burke, 1984). It is however still possible that SS mRNA is present in stem and root tissue in very low levels. Stem tissue was expected to show some SS mRNA. Coruzzi et al. (1984) have shown by Northern-analysis that PS2.4 is expressed in stem tissue. The level of SS gene expression in stems found in this work is therefore surprisingly low. The level of SS mRNA in dark-grown pea apices is considerably lower than that found in light-grown leaves. The major protected fragment from dark-grown leaf RNA corresponds to message from SS47, SS3.6 and rbcS 3A. The band corresponding to PS2.4 (Figure 24, label c) is missing from RNA isolated from dark-grown plants indicating the under-representation of this member of the family in the dark. The lack of PS2.4 in dark-grown tissue may reflect the tight light-regulation of this member of the family or the leaky regulation of SS47 (or SS3.6, or rbcS 3A). No band corresponding to SS8.0 can be seen in RNA from dark-grown leaves; this is not surprising since its low level of expression in the light would prevent its detection in the dark if its expression were reduced by the same amount as SS47.

A protected fragment ~8 b longer than the one

corresponding to SS47 is present in RNA from dark-grown plants. This fragment is also present in light-grown plants but at a lower level (Figure 24, lane 4 and Figure 25, lanes 2 and 3), it is also present in transgenic tobacco plants expressing SS47 at a high level (Figure 41, lanes 6, 7 and 8). The size increase of ~8 b corresponds to an increase in homology of the probe to the RNA by ~8 b; this could obviously occur at either the 5'- or 3'-end of the probe. The 5'-end of the probe is ordinarily reduced by 79 b (Figure 21), so an increase in homology of 8 b would infer homology between the coding region of the gene and the M13 polylinker sequence; this does not occur. An addition of ~8 b of homology at the 3'-end of the probe would correspond to no bases being removed from this end of the probe by S1 nuclease; the probe extends only 8 b beyond the designated cap site. It would seem that the longer protected fragment is the product of base removal only at the 5'-end of the probe. The origins of the extra band could be two-fold. Firstly, the S1 nuclease is limiting and the 8 b at the 3'-end of the probe are not removed; secondly, the presence of an alternative transcription initiation site for SS47 may be the cause. If the S1 nuclease concentration is limiting, a ladder of bands would be expected between the two bands; this is not seen in either Figure 25, lane 6 or Figure 41, lanes 6, 7 and 8. Also, if the S1 nuclease concentration was limiting, the relative level of the extra protected band would not increase in RNA from dark-grown tissue as compared to RNA from light-grown tissue. If an alternative transcriptional start site is responsible for the extra band, this site is used only infrequently in the light,

corresponding to SS47 is present in RNA from dark-grown plants. This fragment is also present in light-grown plants but at a lower level (Figure 24, lane 4 and Figure 25, lanes 2 and 3), it is also present in transgenic tobacco plants expressing SS47 at a high level (Figure 41, lanes 6, 7 and 8). The size increase of ~8 b corresponds to an increase in homology of the probe to the RNA by ~8 b; this could obviously occur at either the 5'- or 3'-end of the probe. The 5'-end of the probe is ordinarily reduced by 79 b (Figure 21), so an increase in homology of 8 b would infer homology between the coding region of the gene and the M13 polylinker sequence; this does not occur. An addition of ~8 b of homology at the 3'-end of the probe would correspond to no bases being removed from this end of the probe by S1 nuclease: the probe extends only 8 b beyond the designated cap site. It would seem that the longer protected fragment is the product of base removal only at the 5'-end of the probe. The origins of the extra band could be two-fold. Firstly, the S1 nuclease is limiting and the 8 b at the 3'-end of the probe are not removed; secondly, the presence of an alternative transcription initiation site for SS47 may be the cause. If the S1 nuclease concentration is limiting, a ladder of bands would be expected between the two bands; this is not seen in either Figure 25, lane 6 or Figure 41, lanes 6, 7 and 8. Also, if the S1 nuclease concentration was limiting, the relative level of the extra protected band would not increase in RNA from dark-grown tissue as compared to RNA from light-grown tissue. If an alternative transcriptional start site is responsible for the extra band, this site is used only infrequently in the light,

hence the low levels of mRNA corresponding to it, but it appears to be used more frequently in dark-grown tissue. Since the S1 probe extends only 8 b beyond the designated transcription start site, a longer probe is required to investigate the length of this postulated message and verify its existence or confirm its artefactual nature.

S1 nuclease analysis has been applied to the study of the expression of SS genes by Coruzzi et al. (1984); the observed protected fragments in their work indicated that the 5'-probe isolated from PS2.4 will not distinguish between SS47, rbcS 3A (Fluhr et al., 1986), SS3.6 (Cashmore, 1983) and SS8.0 (Timko et al., 1985a) (see also Figure 20). The use of a combination of 3'- and 5'-S1 probes from SS47 enables the identification of different members of the gene family from S1 protected fragments and the relative levels of expression of each member to be investigated.

The relative levels of expression of the different SS genes from pea can be investigated by comparison of the levels of the S1 protected fragments from 5'- and 3'-S1 analysis. From Figure 24 and Figure 25, it can be seen that the largest protected fragment (band a in Figure 24) is the major band. This band, as discussed above, corresponds to SS47, SS3.6 (Cashmore, 1983) and rbcS 3A (Fluhr et al., 1986); these three genes therefore account for over 50% of the SS RNA. If however SS47 and rbcS 3A are the same gene, SS47 and SS3.6 account for over 50% of the SS RNA. Band b in Figure 24 corresponds to SS8.0 and is expressed at a low level (~5% of total SS RNA). Band c in Figure 24 corresponds to PS2.4 and accounts for ~40% of the SS RNA.

Due to the possible presence of other ,as yet,

uncharacterised members of the SS multi-gene family, and SSU1, which have not been included in the above considerations, it is possible that the levels indicated above are over estimates. The level of expression of PS2.4 was estimated as 30-35% by Coruzzi et al. (1984); since their S1 probe will distinguish PS2.4 from all other characterised members of the SS family, it is assumed that the RNA giving rise to band c in Figure 24, that accounts for ~40% of the SS RNA, is due to more than one member of the family.

From the 3'-S1 analysis of the SS genes (Figure 25), and the comparison of the 3'-untranslated regions of all characterised pea SS genes (Figure 1), it can be concluded that the doublet at ~230 b is probably due only to SS47 RNA; this doublet accounts for ~60% of the SS RNA. The possibility does however still exist that another gene with an identical 3'-untranslated region to SS47 is present. Table 5 indicates the expected size of S1 protected fragments based on the sequence comparisons made in Figure 1. Band a (~158 b) in Figure 27 corresponds to SS3.6 (Cashmore, 1983), PS2.4 (Coruzzi et al., 1984) and SS8.0 (Timko et al., 1985a) and accounts for ~30% of the SS RNA. Since the 5'-S1 band due in part to SS47 is the major band, and the 3'-S1 band due entirely to SS47 is also the major protected band, the amount of SS3.6 RNA present in the 5'-S1 band must be very low. This assumption is also based on the joint levels of expression of SS3.6, SS8.0 and PS2.4 (Figure 24); PS2.4 accounts for 30-35% of the SS RNA (Coruzzi et al., 1984) and SS8.0 accounts for ~5% of SS RNA (as discussed above). These data taken collectively infer that

SS3.6 accounts for only a few percent of SS RNA. The band labelled b in Figure 27 (~88 b) is due to RNA from SSU1; this accounts for ~10% of the SS RNA. SSU1 RNA may account for the extra RNA that gave rise to band c in Figure 24; this would account for the difference of ~10% between the level of expression of PS2.4 (30-35%) as ascertained by Coruzzi et al. (1984), and the level seen in band c of Figure 24.

The extreme divergence between the 3'-untranslated region of SSU1 and the other characterised members of the SS multi-gene family may indicate a high level of divergence between its 5'-untranslated region and the corresponding region of other SS genes. If the divergence between the 5'-untranslated region of SSU1 and the S1 probe used in this work is high, then the expected S1 protected fragment will be due to homologies in the coding region only and will give rise to a fragment of ~179 b, the same as for PS2.4. The band labelled c in Figure 27 is ~65 b long. This length is less than the length of SS coding region (82 b) present in the probe. The presence of this ~65 b band is due to either an expressed member of the SS multi-gene family that is divergent in its 3'-coding region from SS47, or is due to an artefactual S1 product. The presence of several other bands in lanes 2 and 3 of Figure 27 is most likely due to the break-down product of the S1 probe due to its very high specific activity. These bands can be seen in the lanes labelled P in Figures 26 and 27 and are caused by the decay of the [^{32}P] used to label the probe and subsequent strand breakage. It is possible that some of these bands may correspond to members of the SS multi-gene family that are

expressed at a very low level. There are at least 5 expressed SS genes and possibly more that share homologies to the characterised ones in the 5'- and 3'-untranslated regions; other weakly expressed SS genes may also be present.

It may be concluded from the 5'- and 3'-S1 analysis data that SS47 is the most highly expressed member of the SS multi gene family, giving rise to about 50% of the SS RNA. The other members of the family give rise to lower amounts of SS RNA varying from a few percent (SS3.6) to 30% (PS2.4). The estimated levels may be over-estimates due to other unknown SS genes. The levels of expression of the different SS genes may differ between var. Feltham First and var. Progress No.9; S1 analysis on RNA from var. Progress No.9 with the probe used in these studies would answer this question. The SS genes appear to be transcribed at a low level in the dark, although PS2.4 levels are lower than expected in the dark; 3'-S1 analysis of RNA from dark-grown leaf tissue is needed to confirm this observation. The possible presence of RNA from other SS genes could be investigated by the use of an end-labelled 3'-specific probe; this would resolve the problem of degraded probe fragments that occur with the uniformly labelled probes used in this work. It will not be possibe to accurately assess the levels of expression of the various members of the SS multi-gene family until all expressed members are identified.

From the 5'-S1 analysis of AB RNA in Figure 26, it can be seen that the S1 probe described in Figure 23 does not distinguish between different members of the AB

multi-gene family. The presence of only one size class of protected fragment indicates that either the probe does not hybridise to all members of the family and only identifies some of them or, there is only one subfamily and all members contain very closely related 5'-untranslated regions. By comparison of AB80 and AB66 (Timko et al., 1985a) it can be seen that these two pea AB genes share identical 5'-untranslated regions. From the absence of any other smaller protected fragments, corresponding to homology between the coding region of AB13 and RNA from other AB genes it is concluded that the AB gene family shows remarkable homogeneity between its members. It is very surprising that the predicted size for the protected fragment, of 447 b is the only size fragment present since by comparison of AB13 to both AB80 and AB66 (Timko et al., 1985a) a fragment of 399 b was expected; this extra protected fragment would correspond to AB80 and AB66 RNA. The expected size difference between AB13 RNA and RNA from AB80 and AB66 is due to the variation between the 5'-coding and 5'-untranslated regions of these genes (Figure 19). The expected size difference accounts for an extra 48 b being removed from the 3'-end of the probe where the two sequences diverge. There is no trace of a 399 b protected fragment: this indicates the absence of equivalents to AB80 and AB66 in P.sativum var. Feltham First. Both AB80 and AB66 were isolated from P.sativum var. Progress No.9 (Timko et al., 1985a). This possible intervarietal difference is very surprising indeed since there appear to be very little differences in the SS multi-gene families from these two varieties. It would be interesting to repeat the 5'-S1

analysis with the AB13 probe on RNA isolated from var. Progress No.8; this would indicate whether an equivalent of AB13 occurs in var. Progress No.9. In addition S1 analysis on RNA from var. Feltham First with a 3'-specific probe from AB13 would show whether the 3'-untranslated region of the AB genes from var. Feltham First are as highly conserved as the 5'-untranslated region appears to be.

5. TRANSFER OF GENOMIC SEQUENCES INTO AGROBACTERIUM BINARY VECTORS AND PLANT TRANSFORMATION

5.A Introduction

The vector system used for the transformation of N.tabacum with SS47 and AB13 was the binary vector system of Bevan (1984). This system is based around the plasmid Bin19. This plasmid is capable of replication in both E.coli and A.tumefaciens. Bin19 contains a defective T-DNA region, one deleted for the tumor-inducing genes, yet retains the necessary terminal repeats for transfer into the plant chromosome. Transfer is mediated in a binary vector system by trans-acting factors from a vir region present on a second replicon. The gene to be transferred to the plant chromosome is cloned into the polylinker region of Bin19, recombinants selected on medium containing -galactosidase in E.coli and transferred into A.tumefaciens strain LBA4404 (Hoekema et al., 1983). A.tumefaciens strain LBA4404 was derived from strain Ach5 and harbours a T-DNA deletion derivative of the octopine Ti plasmid pTiAch5 (pAL4404) (de Frammond et al., 1983). The vir region of pAL4404 provides the necessary signals to instigate transfer of the T-DNA region of Bin19, and any DNA within this region, from the A.tumefaciens cell to the plant cell. A.tumefaciens strain LBA4404, E.coli strain HB101 harbouring pRK2013 (Ditta et al., 1980) and E.coli strain DH1 harbouring Bin19 were gifts from Dr.M.Bevan.

The first stage in the introduction of SS47 and AB13, isolated from P.sativum, into N.tabacum was their

transfer from pACYC184 into Bin19. The 7.5 kb EcoRI fragment from pSS47E was chosen for transfer into Bin19 as was the 4.5 kb EcoRI fragment from pAB13E. pSS47E was chosen because it contains an extra 500 bp at the 5'-end of the gene compared to pSS47B (see Section III.2). pAB13E was chosen in preference to pAB13B because it contains the entire AB13 gene whereas pAB13B contains only half of the gene (see Section III.2).

The Bin19 recombinant plasmids were designated Bin19SS47E and Bin19AB13E. The transfer of SS47 and AB13 into Bin19 was performed in vitro and the resulting recombinant plasmids transformed into E.coli. The recombinant binary vectors were mobilised into A.tumefaciens during a tri-parental mating of E.coli strain MC1022 harbouring recombinant Bin19 molecules, E.coli strain HB101 harbouring pRK2013 and A.tumefaciens strain LBA4404, as described in Section II.2.I.i. The resulting mixed culture was grown under antibiotic selection such that only A.tumefaciens harbouring Bin19 would grow (Bevan, 1984). The selection for Bin19 in E.coli is kanamycin at 25 µg/ml whereas A.tumefaciens harbouring Bin19 can survive at 50 µg/ml kanamycin, a level that proves fatal to E.coli. By repeatedly selecting kanamycin-resistant colonies, clones of A.tumefaciens carrying recombinant Bin19 plasmid DNA were isolated. Repeated selection of colonies also ensured that the kanamycin resistance was due to Bin19 and not pRK2013 which is only transiently maintained in A.tumefaciens. If the chromosomal streptomycin resistance marker of A.tumefaciens strain LBA4404 was also used during selection, the growth rate of the agrobacteria was severly reduced.

Having isolated clones of A.tumefaciens harbouring the recombinant Bin19 molecules, it was necessary to confirm that the plasmids had not undergone any rearrangements. A.tumefaciens strain LBA4404 is not recA$^-$; indeed, some of the plant transformation vector systems, for example those of De Block et al. (1983) and Horsch et al. (1984), rely on recombination in vivo between the Ti plasmid and pBR322 to transfer the gene destined for the plant chromosome from pBR322 into the T-DNA of a Ti plasmid carrying both T-DNA and vir regions. The confirmation of recombinant Bin19 plasmid structure was performed by isolation of Bin19 plasmid DNA from A.tumefaciens by a small scale preparation method (Section II.2.B.ii) and by the isolation of total bacterial DNA, followed by cleavage with the appropriate restriction endonuclease to isolate the insert from the Bin19 plasmid, and Southern blotting to identify the insert fragment from the total agrobacterium DNA. Once the structure of the recombinant Bin19 molecules had been confirmed, the agrobacterium clones were used for the transfer of the SS47 and AB13 into the chromosomes of N.tabacum cells.

A short period of time was spent at the Plant Breeding Institute, Cambridge learning the techniques of plant transformation and tissue culture. Tobacco leaf-discs were infected with A.tumefaciens harbouring either Bin19SS47E or Bin19AB13E at the Plant Breeding Institute, Cambridge. Tobacco callus tissue was transferred to Warwick, cultured under kanamycin selection and shoots induced (Bevan, 1984). Due to a growth room malfunction, all this transformed plant tissue was destroyed and the

transformations were repeated at Warwick following the rapid protocol of Horsch et al. (1985). Transgenic plants were regenerated under kanamycin selection and screened for the expression of the introduced genes (Section III.6).

5.B Analysis of recombinant plasmids in E.coli

The genes SS47 and AB13 were transferred from pACYC184 into Bin19 as follows. Plasmid DNA, either pSS47E or pAB13E (1 µg), was cleaved with EcoRI and extracted with phenol (Section II.2.D.i). The DNA was ethanol-precipitated and resuspended in SDW at 50 ng/µl. Bin19 DNA (1 µg) was cleaved with EcoRI (Section II.2.D.i) and dephosphorylated with CIP (Section II.2.D.ii); this DNA was also resuspended at 50 ng/µl. Ligation reactions were carried out as described in Section II.2.D.iii with 100 ng dephosphorylated Bin19 DNA and 50 ng cleaved pSS47E or pAB13E DNA. DNA ligation products were transformed into E.coli strain MC1022 and Bin19 recombinant plasmids selected on 25 µg/ml kanamycin as well as X-gal (Section II.2.H.i). The double selection ensured that only Bin19-containing bacteria grew and recombinant Bin19-containing colonies could be identified by the X-gal colour selection. Small scale plasmid preparations (Section II.2.B.ii) were performed on 10 white colonies from each ligation. The plasmid DNA was cleaved with EcoRI and the products analysed on a 0.7% 1x TAE agarose gel using purified SS47E and AB13E insert DNA as markers (data not shown). Representative bacteria harbouring Bin19SS47E or Bin19AB13E recombinants were cultured and large scale plasmid preparations performed on each (Section II.2.B.i) Plasmid DNA from pSS47E, pAB13E, Bin19SS47E and

Bin19AB13E was cleaved with EcoRI and the products analysed on a 0.7% 1x TAE agarose gel with HindIII-cleaved λDNA as size markers. A comparison of pSS47E with Bin19SS47E is shown in Figure 28 and a comparison of pAB13E with Bin19AB13E is shown in Figure 29. These agarose gels were Southern blotted (Section II.2.4.i.b) and probed with nick-translated insert (Section II.2.F.i.a) from pSSU160 and pFa/b31 respectivly. The resulting autoradiographs are shown in Figures 30 and 31. These data indicate that the desired insert fragments have been transferred from the E.coli plasmid vectors into Bin19.

5.C Triparental matings between E.coli and A.tumefaciens

The recombinant binary vectors Bin19SS47E and Bin19AB13E were transferred from E.coli strain MC1022 to A.tumefaciens strain LBA4404 with the aid of pRK2013. The matings were performed as described in Section II.2.I.i. The overnight liquid culture required for the mating were grown under antibiotic selection. The A.tumefaciens culture was slow growing as compared to the E.coli cultures and appeared very clumpy and stringy. The overnight culture of the tri-parental mating produced confluent bacterial lawns on nutrient-agar plates. After 3 days growth at 30°C on the first selection (50 μg/ml kanamycin), vigorous colonies were present. The colony morphology was characteristic of A.tumefaciens rather than E.coli. Repeated selective streaking of single colonies onto fresh antibiotic plates produced isolated single white colonies after 3 days growth. After three rounds of selection, a colony harbouring Bin19SS47E and one harbouring Bin19AB13E was chosen and


1 2 kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 28. Restriction analysis of pSS47E and Bin19SS47E

Plasmid DNA from pSS47E (1 μg, lane 1) and Bin19SS47E (200 ng, lane 2) was cleaved with EcoRI and the cleavage products fractionated on a 0.7% 1x TAE agarose gel run at 1.5 V/cm for 16 h. Following electrophoresis the gel was stained in 0.1 mg/ml EtBr for 15 min. The sizes indicated in kb are from λDNA cleaved with HindIII. The bands in lane 2 (10 kb and 7.5 kb) are very faint.

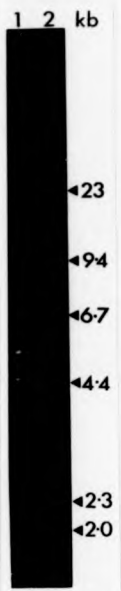

1 2 kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 29. Restriction analysis of pAB13E and Bin19AB13E

Plasmid DNA from pAB13E (1 μg, lane 1) and Bin19AB13E (200 ng, lane 2) was cleaved with EcoRI and the cleavage products fractionated on a 0.7% 1x TAE agarose gel run at 1.5 V/cm for 16 h. Following electrophoresis the gel was stained in 0.1 mg/ml EtBr for 15 min. The sizes indicated in kb are from λDNA cleaved with HindIII.

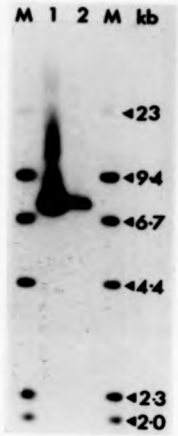

M
1
2
M
kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 30. Southern analysis of restriction endonuclease-digested pSS47E and Bin19SS47E

The DNA fractionated on the gel shown in Figure 28 was transferred to nitrocellulose by capillary blotting; the blot was probed with nick-translated cDNA insert from pSSU160 for 16 h. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat X-ray film with an intensifying screen. HindIII-cleaved λDNA, end-labelled with [α-^{32}P]dGTP was used as size markers (lanes M); the sizes of the fragments are shown in kb. The smear above the band in lane 1 is due to sample overloading on the gel.

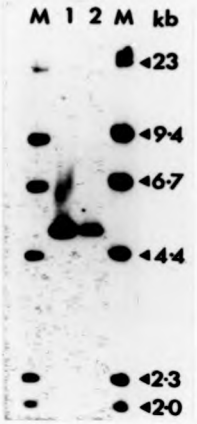

M 1 2 M kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 31. Southern analysis of restriction endonuclease-digested pAB13E and Bin19AB13E

The DNA fractionated on the gel shown in Figure 29 was transferred to nitrocellulose by capillary blotting; the blot was probed with nick-translated cDNA insert from pFa/b31 for 16 h. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat X-ray film with an intensifying screen. HindIII-cleaved λDNA, end-labelled with [α-^{32}P]dGTP was used as size markers (lanes M); the sizes of the fragments are shown in kb.

cultured to analyse the Bin19 recombinants present.

5.D Analysis of recombinant plasmids in A.tumefaciens

i) Plasmid isolation

Bin19SS47E and Bin19AB13E plasmid DNA was isolated from A.tumefaciens as described in Section II.2.B.ii. This DNA (~200 ng) was cleaved with EcoRI and analysed on a 0.7% 1x TAE agarose gel with CsCl-purified, EcoRI-cleaved Bin19 recombinant plasmids (isolated from E.coli) as markers. Comparison of the fragment sizes from the two preparations of plasmid DNA indicated whether the Bin19 recombinants had undergone any rearrangements during their transfer from E.coli to A.tumefaciens. Figure 32 shows the result of this analysis. It can be seen that BinSS47E from A.tumefaciens produces the expected EcoRI fragments whereas Bin19AB13E produces an unexpectedly large insert band. It was assumed that a recombination event had occured to increase the size of the AB13 EcoRI fragment from 4.5 to ~10 kb. A second tri-parental mating was set up to transfer Bin19AB13E into A.tumefaciens and the plasmid isolation repeated. The results from EcoRI cleavage of the second isolate of Bin19AB13E are shown in Figure 33 with linearised Bin19, EcoRI-cleaved pAB13E and HindIII-cleaved λDNA as size markers. This Figure indicates that the second isolate of Bin19B13E has not undergone any rearrangement.

ii) Total DNA isolation

To further confirm the stability of Bin19 recombinants in A.tumefaciens, total DNA from agrobacteria harbouring Bin19SS47E, agrobacteria harbouring first and

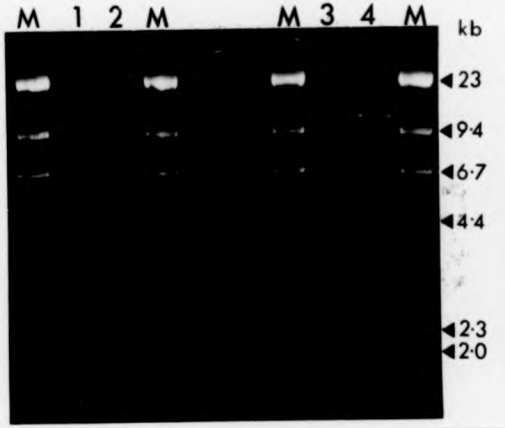

M 1 2 M M 3 4 M
kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 32. Restriction endonuclease analysis of Bin19 recombinant plasmids from A.tumefaciens

Bin19SS47E DNA (100 ng) isolated from A.tumefaciens was digested with EcoRI (lane 1) and Bin19SS47E DNA (500 ng) isolated from E.coli was digested with EcoRI (lane 2). Bin19AB13E DNA (100 ng) isolated from A.tumefaciens (first isolate) was digested with EcoRI (lane 3) and BinAB13E DNA (500 ng) isolated from E.coli was digested with EcoRI (lane 4). The restriction digest products were fractionated on a 0.7% 1x TAE agarose gel. λDNA-cleaved with HindIII was run as size markers (lanes M); the sizes shown are in kb. The bands in lane 1 are faint as is the doublet at ~10 kb in lane 3.

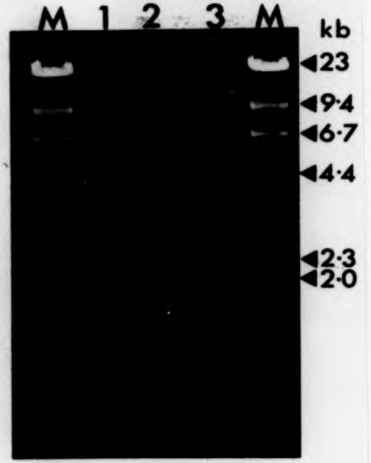
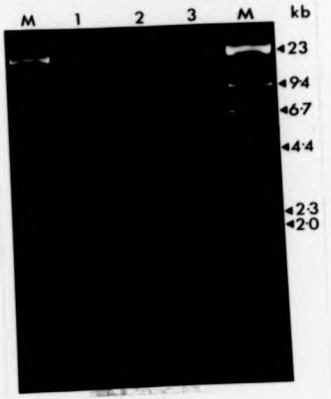

M
1
2
3
M
kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 33. Restriction endonuclease analysis of Bin19AB13E from A.tumefaciens

pAB13E (500 ng), Bin19AB13E (second isolate) (500 ng) and Bin19 (500 ng) DNA was cleaved with EcoRI and fractionated on a 0.7% 1x TAE agarose gel (lane 1, 2 and 3 respectively). λDNA cleaved with HindIII was run as size markers (lanes M); the sizes shown are in kb.


M
1
2
3
M
kb
23
9·4
6·7
4·4
2·3
2·0

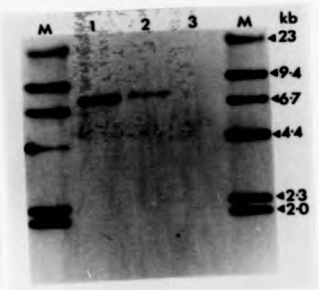
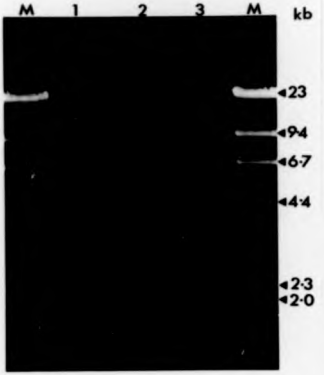

M
1
2
3
M
kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 34. Restriction endonuclease digestion of DNA from A.tumefaciens strain LBA4404 harbouring Bin19SS47E DNA

Total DNA from A.tumefaciens strain LBA4404 (5 μg) was cleaved with EcoRI and the products fractionated on a 0.7% 1x TAE agarose gel (lane 3). Total DNA from A.tumefaciens strain LBA4404 harbouring Bin19SS47E (5 μg) was cleaved with EcoRI and the products fractionated on the same gel (lane 2). Bin19SS47E DNA isolated from E.coli (500 ng) was cleaved with EcoRI and run in lane 1. λDNA cleaved with HindIII was run as size markers (lane M); the sizes are shown in kb.

FIGURE 35. Southern analysis of DNA from A.tumefaciens strain LBA4404 harbouring BinSS47E

The DNA fractionated in the gel shown in Figure 34 was transferred to nitrocellulose and probed with nick-translated cDNA insert from pSSU160. The lanes are as labelled in Figure 34. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat X-ray film with an intensifying screen. HindIII-cleaved λDNA, end-labelled with [α-^{32}P]dGTP, was used as size markers (lane M); the sizes of the fragments are shown in kb.


M
1
2
3
M
kb
23
9·4
6·7
4·4
2·3
2·0

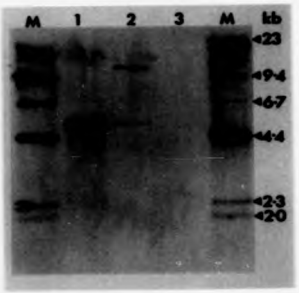

M
1
2
3
M
kb
23
9·4
6·7
4·4
2·3
2·0

FIGURE 36. Restriction endonuclease digestion of DNA from A.tumefaciens strain LBA4404 harbouring Bin19AB13E

Total DNA from A.tumefaciens strain LBA4404 (5 μg) was cleaved with EcoRI and the products fractionated on a 0.7% 1x TAE agarose gel (lane 3). Total DNA from A.tumefaciens strain LBA4404 harbouring Bin19AB13E (1 μg) was cleaved with EcoRI and the products fractionated on the same gel (lane 2). Bin19AB13E DNA isolated from E.coli (500 ng) was cleaved with EcoRI and run in lane 1. λDNA cleaved with HindIII was run as size markers (lanes M); the sizes are shown in kb. Due to the small quantity of DNA loaded in lane 2, the track is very faint.

FIGURE 37. Southern analysis of DNA from A.tumefaciens strain LBA4404 harbouring BinAB13E

The DNA fractionated in the gel shown in Figure 36 was transferred to nitrocellulose and probed with nick-translated cDNA insert from pFa/b31. The lanes are as labelled in Figure 36. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature followed by twice in 0.1x SSC, 0.1% SDS for 15 min at 60°C. The filter was exposed overnight to Kodak X-Omat X-ray film with an intensifying screen. HindIII-cleaved λDNA, end-labelled with [α-^{32}P]dGTP was used as size markers (lane M); the sizes of the fragments are shown in kb.

second isolates of Bin19AB13E and agrobacteria not harbouring Bin19, was isolated (Section II.2.B.vii) and cleaved with EcoRI. The cleavage products were separated on a 0.7% 1x TAE agarose gel (Section II.2.C.i). After being photographed, the gels were blotted onto nitrocellulose (Section II.2.G.i.b) and probed with nick-translated insert (Section II.2.F.i.a) from either pSSU160 or pFa/b31 as appropriate. The DNA isolated from A.tumefaciens harbouring the first isolate of Bin19AB13E (shown to be rearranged in Section III.5.D.i) confirmed the presence of AB13 on an ~10 kb fragment (data not shown). The agarose gel and autoradiograph of DNA isolated from A.tumefaciens harbouring Bin19SS47E are shown in Figures 34 and 35. The gel and autoradiograph of DNA isolated from A.tumefaciens harbouring Bin19AB13E (second isolate) are shown in Figures 36 and 37. EcoRI-cleaved Bin19SS47E or Bin19AB13E and HindIII-cleaved λDNA were used as size markers. End-labelled HindIII-cleaved λDNA was included as markers for the Southern blot. The results confirm that these isolates of A.tumefaciens contain Bin19 recombinants in their desired form.

5.E Plant transformation and tissue culture

Plant transformation techniques were performed as described in Section II.2.I.ii using young leaves from plants of N.tabacum that were approximately 70 days old. During the leaf-disc infection incubation, clumps of agrobacteria could be seen aggregating on the cut leaf edges. Control infections were performed by incubating leaf-discs with A.tumefaciens strain LBA4404 that had not been mated with E.coli harbouring Bin19. Due to the absence

of T-DNA from this culture of agrobacterium, no DNA transfer occurs and kanamycin resistance is therefore not conferred upon the plant cells. During growth of the cultured leaf-discs on selective agarose plates (Section II.2.I.ii) kanamycin-resistant callus tissue forms within 2 weeks. Non-transformed leaf-discs necrose rapidly and after 2 weeks are dead. An example of transformed leaf-discs compared to non-transformed ones is shown in Plate 1. In general, callus tissue develops only at the wounded edges of the leaf-disc but may develop elsewhere on the disc if the surface has been damaged, for example by the forceps used to manipulate the leaf-discs. Non-transformed tissue at the centre of the leaf-disc undergoes necrosis as in the control leaf-discs. The presence of carbenicillin to suppress growth of the agrobacteria was required for several weeks. It was observed that if carbenicillin is omitted from the media after 2 to 3 weeks growth, growth of the agrobacteria resumes. This problem is enhanced if the leaf-discs are not blotted dry after infection with the agrobacteria. To alleviate this problem, the leaf-discs should be thoroughly blotted and carbenicillin retained in the media throughout all tissue culture stages. Selection for transformants was at the level of 300 μg/ml kanamycin. As green kanamycin-resistant callus tissue developed (Plate 2) it was subcultured away from the necrosing leaf-discs onto fresh selective media. The leaf-discs were cultured in 8 cm Petri-dishes as was the initial callus tissue. As shoots developed, the tissue was subcultured into 6 cm high, 4 cm diameter clear Sterilin specimen jars. Root induction was also performed in these jars. Shoots began to appear after 4 weeks and continued to

Plate 1. Transformed and non-transformed tobacco leaf discs

Leaf discs approximately three weeks after incubation with A.tumefacies lacking Bin19 constructs are shown on the left of the picture. Leaf discs approximately three weeks after incubation with A.tumefaciens containing Bin19 constructs are shown on the right of the picture. Selection was at 300 ug/ml kanamycin.

Plate 2. Undifferentiated callus tissue

Undifferentiated callus tissue is shown prior to the formation of shoots. Callus tissue was subcultured every three weeks onto fresh shoot-induction medium. Shoots continued to appear for up to six months.

do so for over 6 months. Dark-green shoots exibiting normal shoot morphology (Plate 3) were excised at the earliest opportunity and placed in rooting medium (Section II.2.I.ii) containing kanamycin at 300 μg/ml.

As well as dark-green shoots exhibiting normal shoot morphology, some other shoot-like structures developed from callus tissue. These shoot-like growths were pale-green and had thick, usually long, leaves; they never rooted but always produced more callus tissue from their base on transfer to rooting medium. Shoots of normal appearance generally rooted within 2 weeks of transfer to rooting medium; however, some turned yellow and died whilst others produced more callus tissue at their base and failed to root. These callus-forming shoots would form roots if all the callus tissue was removed and the shoot transferred to fresh rooting medium. After roots had formed (Plate 4), the plantlets were transferred to potting compost. As much agarose as possible was delicately removed from the roots before transfer to compost. For the first few weeks of growth in compost the plants had very little control of their water balance; high-humidity was maintained around the plants by clear polythene bags. the bags were gradually removed as the plants gained control over their water balance. Without the polythene bags, the transfer from the extremely high humidity of the tissue culture pots to the comparitively low humidity of the growth-room was accompanied by rapid water-loss and subsequent wilting. A transgenic tobacco plant is shown in Plate 5. Twelve transgenic plants containing SS47 and two containing AB13 were regenerated. More AB13-containing plants have since

Plate 3. Shoot-induction from callus tissue

Callus tissue is shown containing shoots after growth on shoot-induction medium. These shoots were excised and transferred to medium to promote root formation.

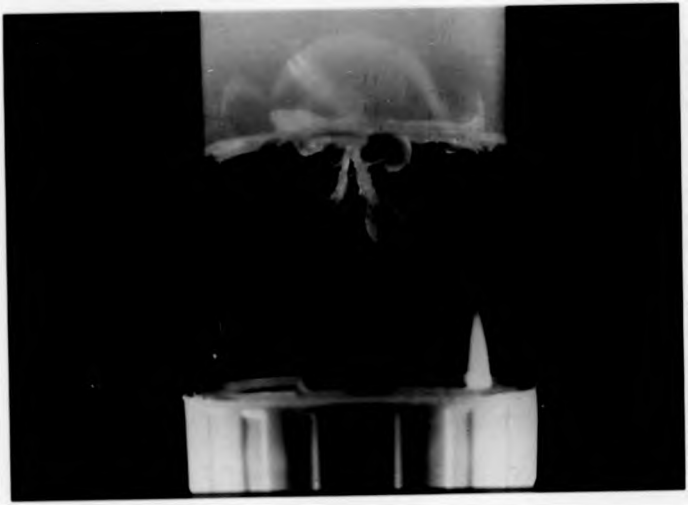


Plate 4. Root-induction from shoots

An excised shoot is shown approximately one week after its transfer to medium containing no phytohormones. The roots are cleary visible.

Plate 5. Regenerated transgenic plant

A regenerated transgenic plant of N. tabacum is shown. It appears normal to the eye, as did all other regenerated plants studied in this work.

been regenerated. All regenerated plants appeared normal to the eye. The transgenic plants were grown to maturity and seeds collected from self-fertilised flowers.

5.F Discussion

From Figures 28 and 29 it can be seen that the DNA fragments containing SS47 and AB13 (7.5 and 4.7 kb respectively) have been transferred into Bin19. Their presence is confirmed by the Southern blots of the gels shown in Figures 28 and 29 (Figures 30 and 31 respectively). The observed rearrangements of Bin19AB13E shown in Figure 32 indicate the need to analyse Bin19 recombinant plasmids once they are transferred into agrobacteria. The Southern blots of total agrobacterium DNA (Figures 35 and 37) indicate that the Bin19 recombinant plasmids Bin19SS47E and Bin19AB13E (second isolate), have not undergone any rearrangements. The larger of the two bands hybridising to pFa/b31 (Figure 37, lane 2) is due to a partial EcoRI digestion product of Bin19AB13E. This partial EcoRI fragment consists of linearised plasmid-insert DNA of ~15 kb.

After 2 weeks growth on kanamycin-selection, transformed and non-transformed tissue can be easily distinguished. Proliferating callus tissue indicates the presence and expression of the kanamycin-resistance gene from Bin19 in the plant chromosome. The roots are the most sensitive organ of the plant to kanamycin since they have the closest contact with it. Shoots that rooted within 2 weeks of transfer to rooting medium are showing kanamycin-resistance to levels of 300 μg/ml; shoots that failed to produce roots and died may be due to the

expression of the kanamycin resistance gene in the shoot at levels too low to enable root formation. Inhibition of root formation by the presence of callus tissue was observed when shoots were excised with some callus tissue attached to the stem. Proliferation of the callus tissue appears to prevent root formation. Subsequent removal of this callus tissue does not further inhibit root formation. The abnormal shoot-like structures that occasionally arose were probably due to either hormone imbalance within the tissue or tissue culture-induced chromosomal rearrangements. This has been reported for potato (see Shepard, 1983), where regenerated plants may exhibit abnormal morphology and chromosomal rearrangements. The inability of plants to control their water balance on removal from tissue culture is not surprising since they have no need of its control in the high humidity environment of a tissue culture pot.

Some of the transgenic plants were screened for the expression of the introduced genes and the results of these experiments are presented in Section III.6. The transmission of the transferred genes to the R_1 generation was investigated and the results of this study are also shown in Section III.6.

6. ANALYSIS OF TRANSGENIC PLANTS

6.A Introduction

The plants regenerated from leaf discs infected with A.tumefaciens carrying SS47 or AB13 sequences were screened for the presence and expression of these genes. The techniques of Southern analysis and S1 nuclease analysis were used to analyse the potential transgenic plants. S1 nuclease analysis of total RNA with pea-specific probes enabled the differentiation between RNA transcribed from the introduced genes and the endogenous genes. As discussed in detail in the literature review at the begining of this thesis, several authors have reported the expression of pea SS genes, or constructs involving the 5'-region of these genes, in transformed tobacco and petunia cells. The levels of expression of pea SS genes in tobacco have been reported as being low (Herrera-Estrella et al., 1984; Nagy et al., 1985), whilst pea SS genes in petunia appear to be expressed at consistently higher levels (Nagy et al., 1985). In contrast to the above reports, the data presented here indicate that high levels of expression of pea SS genes can be obtained in transgenic tobacco plants. The use of S1 nuclease analysis to investigate the expression of introduced genes in transgenic plants has been demonstrated previously (Morelli et al., 1985; Nagy et al., 1985; Schoffl et al., 1985; Goldsborough et al., 1986).

The terminology outlined by Potrykus et al. (1985a) to describe regenerated plants and their progeny has been applied here; regenerated plants are therefore referred to

as R. Selfing of R plants produces the R_1 generation. The transmission of the introduced genes through the germ line to R_1 plants was investigated by S1 nuclease analysis of RNA from selected R_1 plants and by the germination of seeds from transgenic plants under kanamycin selection. The 12 plants regenerated after transformation with SS47 were designated TSS1 to TSS12; those transformed with AB13 were designated TAB1 and TAB2. The R_1 generation plants containing SS47 were designated TSS1-R_1 to TSS10-R_1. TSS9 never produced flowers; TSS11 and TSS12 had not set seed at the time of writing. All R plants appeared phenotypically normal; the only sign of aberrant development observed in the R plants was the failure of TSS9 to flower.

Previous reports of the transmission of introduced genes to the R_1 generation indicate normal Mendelian inheritance with the foreign genes segregating as either one or two independent loci (DeBlock et al., 1984; Potrykus et al., 1985a). The data obtained for all but one of the R_1 generations studied here also indicate normal Mendelian inheritance. However, when the segregation of the kanamycin resistance genes in the R_1 generation from TSS3 was investigated, the ratio of kanamycin-sensitive to kanamycin-resistant plants did not conform to the predicted ratios for normal Mendelian segregation. An explanation for this is proposed and discussed.

6.B Southern analysis of genomic DNA from plants transformed with SS47

Total DNA was isolated from young leaf tissue of plants transformed with SS47 as described in Section

II.2.B.viii. Total DNA was also isolated from non-transformed tobacco leaves and from pea leaves. DNA (10 µg) was cleaved with EcoRI and fractionated on 0.6% 1x TAE agarose gels. EcoRI was chosen since this enzyme should isolate the EcoRI fragment from within the T-DNA and give a hybridisation band of 7.5 kb when the DNA is probed with an SS-specific probe. The DNA was transferred from the agarose gel to either nitrocellulose or Hi-Bond N by capillary blotting (Section II.2.G.i.b) and probed with the EcoRI-SalI fragment from pSS47E (Figure 14) that had been labelled by nick-translation (Section II.2.F.1.a). After exposure to Kodak X-Omat X-ray film for up to two weeks with an intensifying screen, no hybridisation bands could be seen in either the transformed tobacco DNA or the control pea DNA. The experiment was repeated twice more but only weak bands were obtained; these were too faint to be reproduced photographically and are therefore not presented.

The reason for such poor hybridisation results is believed to be the size of the DNA obtained from the mini-preparation method employed, depite the claims of Dellaporta et al. (1983) that DNA of <50 kb could be obtained. CsCl-density gradient centrifugation of the genomic DNA was found to improve the signal strength on Southern blots but by this stage the regenerated plants had set seed and, the only tissue available for DNA extraction was tissue from R_1 generation plants. Since the copy number of SS47 genes in the R_1 generation would be variable as compared to the parent plant, it was decided to concentrate on the expression of SS47 in the regenerated plants. The presence of RNA corresponding to SS47 is proof in itself of

the presence of the gene within the plant cell.

6.C Southern analysis of genomic DNA from plants transformed with AB13

DNA from plants transformed with AB13 was isolated and subjected to Southern-hybridisation as described above in Section III.6.B. The probe used was the entire 4.9 kb insert from pAB13E. As described above, only very weak hybridisation signals were observed; the data are therefore not presented. However all regenerated plants from both SS47 and AB13 transformations were shown to be transgenic by either the demonstration of the presence of RNA from the introduced genes, or by the transmission of kanamycin-resistance through the seed to R_1 plants.

6.D S1 nuclease analysis of total RNA from plants transformed with SS47 using a 5'-specific probe

Plants regenerated from the SS47 transformation were analysed for expression of the introduced gene by S1 analysis using the 5'-specific probe shown in Figure 21. This probe is also species-specific and will differentiate between the RNA transcribed from endogenous tobacco SS genes and that transcribed from the introduced pea SS gene. Annealing reactions and S1 nuclease digestions were performed as described in Section II.2.J.i; conditions for the analysis were as described in Section III.4.B. Total RNA (10 μg) from transgenic plants, and total RNA (10 μg) from non-transformed tobacco leaves was annealed to 50 pg prime-cut probe at 70°C for 3 h. Total RNA (1, 5 and 10 μg) from light-grown pea leaf RNA was also annealed, as

decribed above, to 50 pg prime-cut probe. X.borealis tadpole RNA (10 μg) was annealed to 50 pg prime-cut probe as a control to identify any protected fragments due to probe-probe interactions or hybridisation of the probe to non-SS RNA. The plants used in the S1 analyses are indicated in the appropriate Figure legends. The S1-protected fragments were analysed on 6% polyacrylamide gels and exposed to X-ray film overnight. The results of the 5'-S1 analysis on selected plants are shown in Figures 38 and 39. Figure 39 also shows the analysis of RNA from several R_1 generation plants. These Figures demonstrate the expression of SS47 in transgenic tobacco and indicate the species-specificity of the S1 probe used.

6.E S1 nuclease analysis of total RNA from plants transformed with SS47 using a 3'-specific probe

S1 analysis using the 3'-probe shown in Figure 22 was performed on RNA from a selection of the SS47-transgenic plants. RNA (10 μg) from transformed plants, from non-transformed plants, and from X.borealis was annealed to 50 pg prime-cut probe for 3 h at 70°C (Section II.2.J.i). RNA (1, 5 and 10 μg) from light-grown pea leaf tissue was also annealed to 50 pg of probe. The RNA-DNA hybrids were subjected to S1 nuclease digestion for 30 min at 30°C and the protected fragments fractionated on a 6% polyacrylamide gel (Section II.2.C.iii.a). Figure 40 shows the results of such an experiment. As suggested by Figures 38 and 39, Figure 40 indicates both the expression of SS47 in various transgenic tobacco plants and the species-specificity of the 3'-S1 probe.

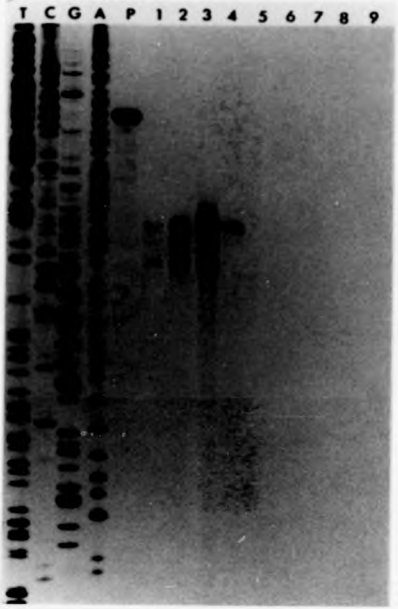

T C G A P 1 2 3 4 5 6 7 8 9

FIGURE 38. 5'-S1 analysis of RNA from transgenic tobacco plants containing SS47

Total RNA (1, 5 and 10 μg) from 9-day-old pea plants grown in the light was annealed to 50 pg of the single-stranded 5'-specific S1 probe shown in Figure 21. Total RNA (10 μg) from transgenic tobacco TSS1, TSS2, TSS3 and TSS4 and from non-transformed tobacco was annealed to 50 pg single-stranded probe, as was 10 μg total RNA from Xenopus borealis (a gift from C.Wilson). Annealing was performed at 70°C for 3 h. RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on a 6% denaturing polyacrylamide gel at 40 W for 1.5 h. The gel was exposed to X-ray film overnight. M13 sequence reactions of the S1 probe were run as size markers. The reactions are labelled according to the ddNTP present. The S1 probe (~5 pg) is shown in the lane labelled P. Lanes 1, 2 and 3 contain the S1-protected fragments from 1, 5 and 10 μg pea RNA respectively. Lanes 4, 5, 6 and 7 contain the S1-protected fragments from TSS1, TSS2, TSS3 and TSS4 RNA respectively. The S1 products from non-transformed tobacco were run in lane 8 and the products from X.borealis RNA were run in lane 9.

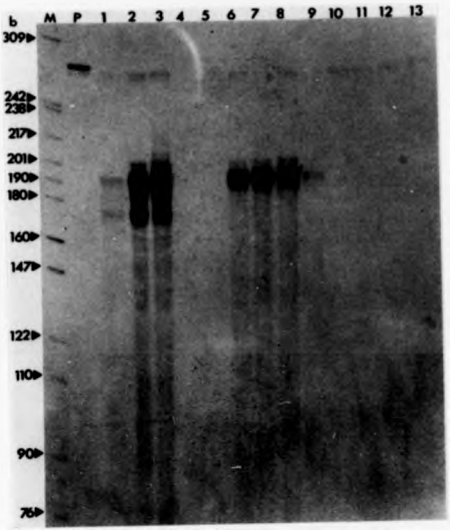

b
M
P
1
2
3
4
5
6
7
8
9
10
11
12
13
309
242
238
217
201
190
180
160
147
122
110
90
76

FIGURE 39. 5'-S1 analysis of RNA from further transgenic tobacco plants containing SS47

Total RNA (1, 5 and 10 μg) from 8-day-old pea plants grown in the light was annealed to 50 pg of the single-stranded 5'-specific S1 probe shown in Figure 21. Total RNA (10 μg) from transgenic tobacco TSS6, TSS7, TSS8, TSS9, TSS10, TSS1-R_1, TSS2-R_1 and TSS4-R_1 and from non-transformed tobacco was annealed to 50 pg single-stranded probe, as was 10 μg total RNA from Xenopus borealis (a gift from C.Wilson). Annealing was performed at 70°C for 3 h. The RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on a 6% denaturing polyacrylamide gel at 40 W for 1.5 h. The gel was exposed to X-ray film overnight. End-labelled HpaII restriction products of pBR322 were run as size markers (a gift from A.Mohamed), the sizes of which are indicated in bases. The S1 probe (~5 pg) is shown in the lane labelled P. Lanes 1, 2 and 3 contain the S1-protected fragments from 1, 5 and 10 μg pea RNA respectively. Lanes 4, 5, 6, 7, 8, 9, 10 and 11 contain the S1-protection products from TSS6, TSS7, TSS8, TSS9, TSS10, TSS1-R_1, TSS2-R_1 and TSS4-R_1 RNA respectively. The S1 products from non-transformed tobacco were run in lane 12 and the products from X.borealis RNA were run in lane 13. The absence of any signal in lane 4 is due to the loss of the sample during annealing.

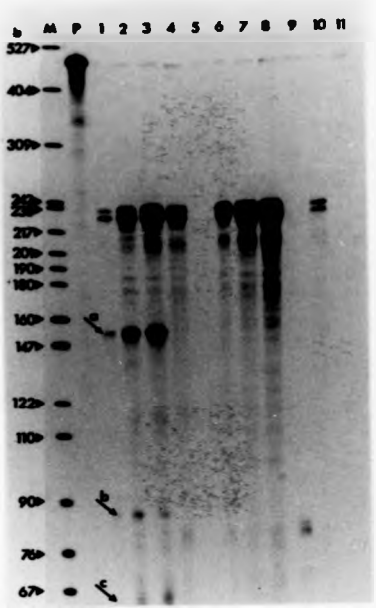

b
M
P
1
2
3
4
5
6
7
8
9
10
11
527
404
309
242
238
217
201
190
180
160
147
122
110
90
76
67
a
b
c

FIGURE 40. 3'-S1 analysis of RNA from transgenic tobacco plants containing SS47

Total RNA (1, 5 and 10 μg) from 9-day-old pea plants grown in the light was annealed to 50 pg of the single-stranded 3'-specific S1 probe shown in Figure 22. Total RNA (10 μg) from transgenic tobacco TSS1, TSS6, TSS8, TSS9, TSS10, TSS1-R_1, and from non-transformed tobacco was annealed to 50 pg single-stranded probe, as was 10 μg total tadpole RNA from Xenopus borealis (a gift from C.Wilson). Annealing was performed at 70°C for 3 h. RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on a 6% denaturing polyacrylamide gel at 40 W for 1.5 h. The gel was exposed to X-ray film overnight. End-labelled HpaII restriction products of pBR322 were run as size markers (a gift from A.Mohamed), the sizes of which are indicated in bases. The S1 probe (~5 pg) is shown in the lane labelled P. Lanes 1, 2 and 3 contain the S1-protected fragments from 1, 5 and 10 μg pea RNA respectively. Lanes 4, 5, 6, 7, 8, 9 and 10 contain the S1-protection products from TSS1, TSS6, TSS8, TSS9, TSS10, non-transformed tobacco and TSS1-R_1 respectively. The products from X.borealis RNA were run in lane 11. The letters a and b indicate bands protected by members of the SS multigene other than SS47. The source of band c was discussed in relation to Figure 27.

6.F S1 nuclease analysis of total RNA from plants transformed with AB13 using a 5'-specific probe

Of the plants transformed with AB13, only TAB1 was screened for the expression of this gene; no other plants were available at that time. The probe used is shown in Figure 23. RNA (10 μg) from TAB1, non-transformed tobacco and X.borealis was annealed to 50 pg of the S1 probe at 70°C for 3 h (Section II.2.J.i). Pea leaf RNA (1, 5 and 10 μg) from light-grown plants was also annealed to 50 pg of S1 probe. The RNA-DNA hybrids were digested with S1 nuclease as described in Section II.2.J.i and the protected fragments analysed on a 6% polyacrylamide gel (Section II.2.C.iii.a). The gel is shown in Figure 41. No AB13 mRNA was detected in TAB1 total RNA.

6.G Transmission of introduced genes through the seed of transformed plants to the R_1 generation

Seeds collected from self-pollinated transgenic plants were surface-sterilised and germinated under kanamycin-selection as described in Section II.2.I.iii. Both kanamycin-resistant and kanamycin-sensitive plants germinated at 100 μg/ml kanamycin; kanamycin-sensitive plants however turned yellow and died after approximately one week, whilst the resistant plants remained bright green and continued to grow. Plate 6 shows the germination of seeds obtained from self-pollination of TSS4 on 100 μg/ml kanamycin; the kanamycin-resistant and kanamycin-sensitive phenotypes can be unambiguously identified. The segregation data obtained were compared to the predicted Mendelian

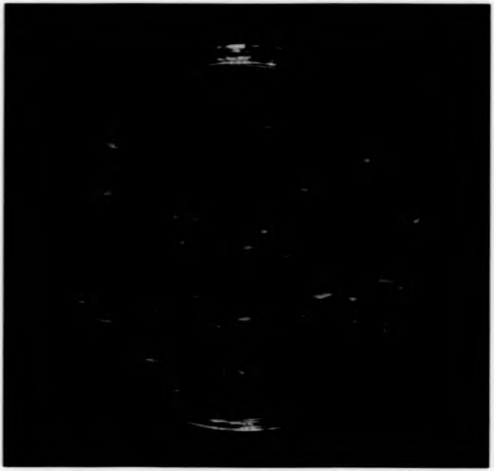


Plate 6. Seeds from a transgenic plant after germination on kanamycin

Seeds collected from selfed flowers of TS54 where surface sterilised and sown on medium containing no phytohormones. Selection was at 100 µg/ml kanamycin. The photograph was taken twelve days after sowing the seeds. The kanamycin-sensitive and kanamycin-resistant plants are clearly distinguishable.

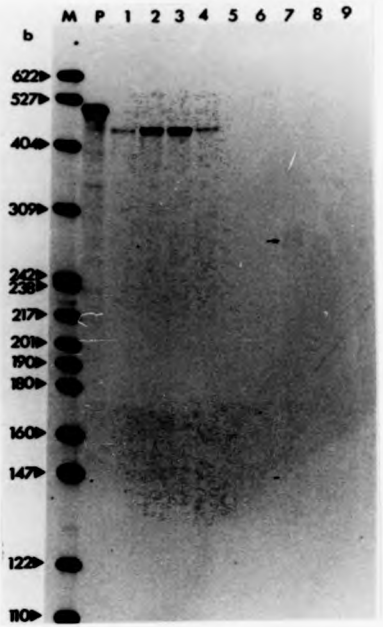

M P 1 2 3 4 5 6 7 8 9
b
622
527
404
309
242
238
217
201
190
180
160
147
122
110

FIGURE 41. 5'-S1 analysis of RNA from transgenic tobacco plants containing AB13

Total RNA (1, 5 and 10 μg) from 9-day-old pea plants grown in the light was annealed to 50 pg of the single-stranded 5'-specific S1 probe shown in Figure 23. Total RNA (15 μg) from transgenic tobacco TAB1 and total RNA (10 μg) from non-transformed tobacco was annealed to 50 pg single-stranded probe, as was 10 μg total RNA from Xenopus borealis (a gift from C.Wilson). Annealing was performed at 70°C for 3 h. RNA-DNA hybrids were digested with S1 nuclease (1000 u/ml) at 30°C for 30 min. The protected fragments were fractionated on a 6% denaturing polyacrylamide gel at 40 W for 1.5 h. The gel was exposed to X-ray film overnight. End-labelled HpaII restriction products of pBR322 were run as size markers (a gift from A.Mohamed), the sizes of which are indicated in bases. The S1 probe (~5 pg) is shown in the lane labelled P. Lanes 1, 2 and 3 contain the S1-protected fragments from 1, 5 and 10 μg pea RNA respectively. The samples run in lanes 4, 5 and 6 are discussed in Figure 26. The S1-nuclease digestion products from TAB1 RNA were run in lane 7 and the products from non-transformed tobacco were run in lane 8. X.borealis tadpole RNA samples were run in lane 9. The arrow indicates a faint protected doublet that is due to cross-reactivity of the pea AB13 probe to tobacco AB RNA.

ratios using the χ^2 test. For Mendelian inheritance, the ratio of kanamycin-sensitive plants to kanamycin-resistant plants in the R_1 generation will be 1:3 for one locus, 1:15 for two independently-segregating loci, 1:63 for three independently-segregating loci and 1:255 for four independently-segregating loci. The formula for this analysis is as follows:

$$\chi^2 = \frac{(O-E)^2}{E}$$

where O is the observed value and E is the expected value.

The germination data, along with the χ^2 values, the segregation ratios and the number of segregating loci are presented in Table 6. The germination frequency was over 95% for seeds from all plants. Non-transformed tobacco seeds were also germinated on 100 μg/ml kanamycin; these seeds were 100% kanamycin-resistant. R_1 generation plants segregate with typical Mendelian ratios for one, two or three independent loci, with the exception of TSS3-R_1 in which the ratios of kanamycin-sensitive and resistant plants deviate from regular Mendelian inheritance.

6.H Discussion

The absence of any data from the Southern analysis of DNA from plants transformed with SS47 and AB13 prevents the calculation of the copy number of these genes in the regenerated plants. Some information concerning the number of loci introduced into the transgenic plants can however be gained from the frequency of kanamycin-resistant R_1 seedlings.

The presence of SS47 RNA in some transgenic plants demonstrates the presence and activity of the SS47 gene in

				χ^2 values					
Plant	Kan^S	Kan^R	Total	1:3	1:15	1:63	1:255	Ratio	Loci
TSS1	3	156	159	---	5.17	0.11	9.17	1:63	3
TSS2	8	168	176	39.3	0.87	10.2	---	1:15	2
TSS3	38	243	281	19.7	25.4	261	---	?	?
TSS4	27	82	109	0.003	63.9	---	---	1:3	1
TSS5	31	88	119	0.07	79.6	---	---	1:3	1
TSS6	11	232	243	54.3	1.23	13.9	---	1:15	2
TSS7	6	288	294	---	8.89	0.44	20.5	1:63	3
TSS8	55	179	234	0.28	118	---	---	1:3	1
TSS10	13	256	269	58.4	0.92	18.7	---	1:15	2
TAB1	89	234	323	1.12	250	---	---	1:3	1

Table 6. Analysis of kanamycin-resistance segregation in the R_1 generation

The regenerated plants used in this study are indicated in the left hand column. The seeds obtained from these plants after self-pollination were germinated under sterile conditions on 100 µg/ml kanamycin. The numbers of kanamycin-sensitive seedlings are indicated in the column labelled Kan^S and the numbers of kanamycin-resistant seedlings are indicated in the column labelled Kan^R. The total number of seeds screened from each plant is indicated in the column labelled Total. The χ^2 values obtained from the figures in the first three columns are shown under the headings of the predicted ratios. The figures underlined are the ones that indicate agreement, at the 5% confidence level, to the predicted ratios. The three dashes (---) indicate where χ^2 values were not calculated for that particular segregation ratio. The segregation ratio confirmed statistically is indicated in the column labelled Ratio and the number of independently-segregating loci inferred by this ratio is indicated in the column labelled Loci. The R_1 generation obtained from TSS3 do not conform to any of the predicted Mendelian segregation ratios; the ratio and number of loci are therefore indicated by question marks (?).

the plants. From Figure 38, lanes 1, 2 and 3, it can be seen that the S1 probe is in excess by the increase in the signal observed as pea RNA quantity is increased. Lane 4 of Figure 38 shows a band due to S1 nuclease protection of the probe by SS47 RNA present in TSS1 total RNA; only the band corresponding to SS47 is present, none of the other bands seen in pea RNA are present. This is as expected since the 7.5 kb EcoRI fragment containing SS47 that was introduced into tobacco cells does not contain any other SS genes. The intensity of the band in Figure 38, (lane 4) protected by 10 μg TSS1 total RNA is slightly greater than the corresponding band from 1 μg of pea RNA. Since the corresponding band in pea is protected by SS3.6 and SS47 RNA whilst in TSS1 it is due entirely to SS47 RNA it can be concluded that the level of SS47 expression in TSS1 is over 10% of the level found in pea. Since SS47 RNA accounts for over 50% of total SS RNA in pea (Section III.4.E), the level of SS47 expression as related to total pea SS expression, is over 5%. This level is higher than that obtained by Nagy et al. (1985) but comparable with the level of expression of SS3.6 in tobacco callus obtained by Herrera-Estrella et al. (1984). Since SS3.6 accounts for only a few percent of pea SS RNA (Section III.4), the comparable levels of expression of SS47 (this work) and SS3.6 (Herrera-Estrella et al., 1984) indicate that SS47 is not expressed in TSS1 at a level comparable to that in pea, whilst SS3.6 is expressed in tobacco callus tissue at a level comparable to its expression in pea. The level of SS3.6 expression in callus tissue observed by Herrera-Estrella et al. (1984) may however be nonrepresentative of the levels that would be obtained in a

regenerated plant because of the nonmorphogenic nature of the callus tissue (Fluhr et al., 1986). RNA from TSS2, TSS3 and TSS4 shows no detectable expression of SS47 (Figure 38, lanes 5, 6 and 7); longer exposure of the gels does not reveal any protected bands. Lane 8 of Figure 38 corresponds to tobacco RNA annealed to the pea S1 probe and digested with S1 nuclease; the absence of protected fragments in this lane indicates the species-specificity of the probe. The absence of any bands in lane 9 of Figure 38 (X.borealis RNA) show that there is no homology to non-SS RNA. Figure 39 shows 5'-S1 analysis on a further selection of transgenic plants. Lanes 1, 2 and 3 of Figure 39 show bands protected by 1, 5 and 10 μg of pea RNA respectively. lane 4 represents 10 μg TSS6 RNA but most of the sample was lost during the RNA-DNA annealing reaction when the capillary tube leaked: what remained of the sample was run on the gel. No conclusions can therefore be drawn from this lane. TSS6 RNA was re-analysed at a later date; the results of this are shown in Figure 40 and discussed below along with the rest of the data from this Figure. Lane 5 of Figure 39 contains 10 μg TSS7 RNA, but no bands are visible; after a longer exposure of 3 days, a very faint band due to SS47 expression was observed (data not shown) indicating the presence of a very small amount of SS47 RNA in TSS7 RNA. Lanes 6, 7 and 8 of Figure 39 show S1-protected bands due to 10 μg RNA from TSS8, TSS9 and TSS10 respectively. As can be seen from the intensity of the bands in lanes 6,7 and 8 of Figure 39, the level of expression of SS47 in the tobacco plants analysed in these lanes is similar to the level found in pea (Figure 39, lane 3).It has been noted by Jones et al. (1985), that

the site of integration of a gene in a foreign chromosome may cause a variation in the level of expression of that gene by up to 200-fold between different transgenic plants. As discussed by Fluhr *et al.* (1985), the position effects observed in transgenic plants and animals may be due to insertion of the gene into heterochromatin, DNA methylation (van Slogteren *et al.*, 1984b), or the presence of vector sequences in close proximity to the gene under study, as found in transgenic mice (Townes *et al.*, 1985). The level of expression of foreign genes in transgenic mice has been shown to be variable and independent of copy number (for review see Palmiter and Brinster, 1985).

The band observed in lane 9 of Figure 39 is due to S1-nuclease protection of the probe by 10 μg RNA isolated from TSS1-R_1 plants. Five TSS1-R_1 plants germinated without kanamycin selection were used for RNA preparation. The band in lane 9 of Figure 39 indicates the transmission of SS47 through the seed and its continued expression in R_1 plants. The intensity of the signal obtained from SS47 in TSS1-R_1 (Figure 39) appears slightly lower than the corresponding signal obtained from TSS1 (Figure 38); this is deduced by comparison of the two signals to those of 1 μg pea RNA on Figures 38 and 39 respectively. This reduction of SS47 RNA in TSS1-R_1 as compared to TSS1 is also seen in the 3'-S1 analysis shown in Figure 40. RNA from TSS2-R_1 and TSS4-R_1 (Figure 39, lanes 10 and 11) show no expression of SS47 as observed for their parent plants.

From the 3'-S1 analysis of SS47-transformed plants (Figure 40), it can be seen that the same doublet observed

in pea RNA at approximately 235 b is present in RNA from transgenic tobacco plants. This observation indicates that SS47 RNA is the source of the two bands since no other pea SS gene is present in the transgenic tobacco plants. The implications of this conclusion are discussed in Section III.4.E. The absence of any protected fragments in lane 5 of Figure 40 indicate the absence of SS47 RNA in RNA isolated from TSS6; longer exposure of the gel does not reveal even a weak signal. The strong signals observed in lanes 6, 7 and 8 of Figure 40 are due to high levels of SS47 RNA in total RNA from TSS8, TSS9 and TSS10 respectively, as observed from 5'-S1 analysis (Figure 39). Lane 9 of Figure 40 contains a faint band at ~80 b, this is due to tobacco SS RNA. The bands other than the 235 b doublet in lanes 1 to 8 and those labelled a, b and c in lanes 1 to 3 are due to artefactual S1 digestion of the RNA-DNA hybrids and S1 probe breakdown.

A comparison of signal strength between lane 4 of Figure 40 (10 µg TSS1 RNA) and lane 10 of Figure 40 (10 µg TSS1-R_1 RNA) indicates that the level of expression of SS47 in the R_1 plants is lower than the level observed in the parent plant. The possible reason for this is the use of R_1 plants that were not germinated under kanamycin selection; after self-fertilisation some of the progeny will contain more SS47 genes than the parent plant, some will contain less. The level of expression of SS47 in the R_1 generation will therefore depend on the number of expressed SS47 genes that individual progeny inherit from the parent plant. The level of expression observed in the five progeny pooled for this analysis may therefore be a biased representation of the average level of expression in the R_1 generation.

Individual R_1 plants need to be tested for their level of expression of SS47 as do individual R_2 plants (obtainable from self-pollinated R_1 plants) to assess the stability of expression of the introduced gene in transgenic plants. The variation in levels of expression of SS47 in various transgenic plants is consistent with previous observations (Jones et al., 1985; Fluhr et al., 1986) of clonal variation. The observations that pea genes in transgenic tobacco are expressed weakly (Fluhr et al., 1986; Herrera-Estrella., 1984) may be a consequence of the fact that the SS genes used by these two groups are not the most abundantly expressed SS genes in pea. The data presented here demonstrate that high levels of expression of pea genes in tobacco can be obtained when the gene under investigation is highly expressed in its parent plant.

By comparison of the S1 products obtained from SS47 in pea and transgenic tobacco RNA, it can be seen that the correct transcription start and stop sites are used in transgenic tobacco. This observation is in agreement with the results obtained from other plant genes introduced into heterologous plant species (Matzke et al., 1984; Morelli et al., 1985; Nagy et al., 1985; Schoffl et al., 1985; Goldsborough et al., 1986)

The expression of AB13 in transgenic tobacco was not evident from the S1 analysis data presented in Figure 41. Lanes 4, 5 and 6 are as described in Figure 26. Lane 7 represents 15 μg TAB1 RNA and lane 8, 10 μg tobacco RNA. A longer exposure of 4 days does not reveal any protection of the S1 probe by TAB1 RNA. TAB2 tissue was not available at the time the S1 analyses were performed and is as yet

uncharacterised. AB13-transformed callus tissue is still in tissue culture and several shoots are present. These shoots will be regenerated into plants and screened for AB13 expression at a later date. The absence of expression of AB13 in TAB1 may be due to either integration into an unfavourable site, as presumed for TSS2, TSS3, TSS4 and TSS6, or that AB13 is a weakly expressed member of the multi-gene family. It is not possible to draw conclusions from these preliminary data. Further work is needed to characterise other plants transformed with AB13 before comparisons can be made to the work of other researchers.

The data obtained from the germination of seeds from transformed plants on kanamycin, summarised in Table 6, enabled the calculation of the number of independently-segregating kanamycin-resistance loci in the transformed plants. The observed 1, 2 or 3 independently-segregating loci are consistent with conclusions obtained by several other workers (De Block *et al.*, 1984; Horsch *et al.*, 1984; Hain *et al.*, 1985; Potrykus *et al.*, 1985a). The non-Mendelian segregation of kanamycin-resistance loci in seeds from TSS3 is discussed below.

The presence of one kanamycin-resistance locus does not necessarily mean that only one T-DNA insertion event has occurred. It has been reported on several occasions that T-DNA insertion may occur as a tandem repeat of T-DNA sequences (Lemmers *et al.*, 1980; Ooms *et al.*, 1982; Holsters *et al.*, 1983). If such tandem insertion events have occurred in the transformed plants studied here, then it is not possible to infer the copy-number of T-DNA sequences inserted from the number of independently-segregating loci.

However, it is possible to say that there are at least as many T-DNA copies per cell as there are independently-segregating kanamycin-resistance loci. Since the kanamycin-resistance marker and SS47 are physically linked, it is assumed that the copy-number of SS47 per cell is at least as many as the number of kanamycin resistance loci. Data from Southern blots would be valuable for the determination of gene copy number, rearrangements and tandem or single insertion events. From a consideration of the number of introduced loci, it is possible to assess the effect of copy number on the level of SS47 expression. These data are summarised in Table 7. It can be concluded that the level of expression of the introduced gene bears absolutely no relation to the copy-number of the gene. The site of insertion of the T-DNA is therefore presumed to be the single most important factor affecting the expression of transferred genes. Also, the ability of the chimaeric nopaline synthase-kanamycin resistance gene to function in regions of DNA where SS47 is unable to do so, indicates a major difference in the function and organisation of the constitutive nopaline synthase promoter as compared to the light-inducible and organ-specific SS promoter. These observations on gene expression in relation to copy-number are consistent with the observations from other transgenic plants (Jones et al., 1985; Fluhr et al., 1986) and transgenic mice (Palmiter and Brinster, 1985). The implications of these observations for the genetic engineering of plant genomes is that heterologous plant promoters can give high levels of expression in transgenic plants if the site of T-DNA insertion is favourable.

Plant	5'-S1 data	3'-S1 data	Loci
TSS1	++ (38)	++ (40)	3
TSS1-R_1	++ (39)	++ (40)	nd
TSS2	- (38)	nd	2
TSS2-R_1	- (39)	nd	nd
TSS3	- (38)	nd	2 (?)
TSS4	- (38)	nd	1
TSS4-R_1	- (39)	nd	nd
TSS5	nd	nd	1
TSS6	nd	- (40)	2
TSS7	+ (39)	nd	3
TSS8	+++ (39)	+++ (40)	1
TSS9	+++ (39)	+++ (40)	nd
TSS10	+++ (39)	+++ (40)	2
TAB1	- (41)	nd	1

Table 7. Summary of expression data from transgenic plants

The transgenic plants studied are listed in the first column. The level of expression of the introduced gene, either SS47 or AB13, in the transgenic plant is indicated by the symbols +, ++, +++ and -. + indicates very low levels of expression, ++ indicates expression at ~10% of the level for that gene in pea, +++ indicates a level of expression approximately equal to the level for that gene in pea. The letters nd denote no available data. The expression of the introduced genes was investigated by either 5'- or 3'-S1 analysis; the source of the data shown is indicated. The Figure from which the data were obtained is indicated in brackets. The number of independently-segregating loci giving rise to the observed expression is indicated in the final column. The question mark (?) referring to the number of loci present in TSS3 indicates that the the number shown is deduced from an aberrant R_1 segregation ratio (see text).

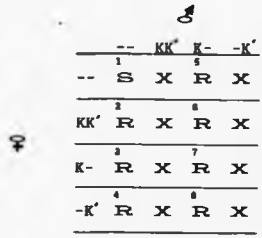
The observation that the kanamycin-resistant phenotype segregates with a ratio of ~1:7, i.e. 38 kanamycin-sensitive to 243 kanamycin-resistant seedlings, in the R_1 generation, and not with any of the predicted Mendelian ratios, indicates that TSS3 may be of special interest. Possible reasons for this observation are linkage of the T-DNA inserts (46%), T-DNA instability leading to a reduction in the number of kanamycin-resistant loci, or that the seeds collected and bulked from all flowers of TSS3 may be the product of a chimaeric plant composed of tissue with a variable T-DNA copy-number.

One more interesting explanation that I should like to propose is the insertional inactivation, by the T-DNA, of a gene required specifically in haploid tissue. The ratio of 1:7 suggests a 1:15 ratio that is lacking an entire class of progeny. It is proposed that one copy of the T-DNA has inserted into a gene essential to the viability of either the pollen or the ova. The gametes that inherit the inactivated gene, yet require its expression, will not be viable; those gametes that inherit the inactivated gene, but do not require its expression, will not be affected. The progeny that the inviable gametes would have produced are therefore absent, resulting in perturbation of the expected ratios. The χ^2 test, when applied to the ratio of 38:243 shows this ratio to be consistent with a ratio of 1:7; it is however also consistent with ratios of 1:6 and 1:8 (sensitive to resistant plants). These ratios can be predicted from two insertional inactivation events, although this is most unlikely. If both inactivated genes, when present in any haploid tissue, are lethal then a 1:6 ratio

would result. If either inactivated gene when present in homozygous form is lethal, then a ratio of 1:8 would result. These two possible explanations are however less likely than the one resulting in the 1:7 ratio but cannot be ruled out at present.

To simplify a further explanation of the above hypothesis I will presume that it is the pollen that require the functional gene. If TSS3 contains two kanamycin-resistance loci, as predicted by the disturbed 1:15 ratio, the gametes produced will be of four classes. The usual sixteen progeny genotypes are reduced to eight by the absence of pollen that have inherited the insertionally-inactivated gene. In Table 8, the genotypes of the gametes (indicated by the numbers) and the phenotype of the progeny are summarised. The eight progeny genotypes absent are also indicated. The hypothesis can thus be tested by growing a random selection of R_1 progeny to maturity and self-pollinating them. The segregation of kanamycin-resistance loci in the R_2 generation will confirm or refute this hypothesis. The ratio of resistant to sensitive phenotypes can be predicted for each of the genotypes listed in Table 8. These results will also test the alternative possibilities of T-DNA instability and the chimaeric nature of TSS3. In the unlikely event that two insertional inactivation events have occurred, the distribution of phenotypes in the R_2 generation will reflect this.

By using other flowers on the flower head of each R_1 plant as either pollen-donors or pollen-acceptors with wild-type tobacco plants, the conclusions from the R_2



	R_1 genotypes		R_2 segregation frequency
1	-/-	-/-	0:4
2	-/K	-/K'	7:1
3	-/K	-/-	3:1
4	-/-	-/K'	1:1
5	K/-	-/-	3:1
6	K/K	-/K'	4:0
7	K/K	-/-	4:0
8	K/-	-/K'	7:1

Table 8. Predicted phenotype and genotype of TSS3-R_1 plants

This Table shows the predicted outcome from the self-fertilisation of TSS3. The genotype of the male gametes is shown horizontally, that of the female gametes vertically. As described in the text, the assumption is that pollen carrying the insertionally-inactivated gene are not viable; they do not therefore produce offspring. The symbols K and K' refer to the two independent kanamycin-resistance loci. The locus labelled K' is the one responsible for gamete inviability. The phenotypes of the R_1 plants are indicated: S indicates kanamycin-sensitive, R indicates kanamycin-resistance and X indicates the absence of progeny due to inviable pollen. The numbers refer to the predicted genotypes of these progeny. These are drawn out under the heading R_1 genotypes. The expected segregation frequencies of the kanamycin-resistant to kanamycin-sensitive phenotypes in the R_2 generation are also shown. The R_1 plant that gives rise to a 1:1 segregation in its progeny is the one best suited to isolation of the insertionally-inactivated gene (see text)

germinations can be confirmed. The out-crossing of each R_1 plant would also reveal whether the inactivated gene affected either the pollen or the ova. If pollen is affected, pollen from R_1 plants used to fertilise wild-type flowers will give rise to disturbed kanamycin-resistance segregation in the progeny. Fertilisation of R_1 plants with wild-type pollen will not affect the expected ratios. If the inactivated gene is required by the ova, the reciprocal results will be observed (see Table 9).

If leaf tissue from the various R_1 plants were collected and frozen before the plants flowered and died, once the genotype of each plant had been identified, the tissue could be used to construct a genomic library from which could be isolated the defective gene by using the T-DNA as a gene tag. The isolated inactivated gene could then be used as a probe to isolate the functional gene from a genomic library constructed from wild-type tobacco plants. To facilitate isolation of the inactivated gene, tissue from plants of genotype 4 (see Table 8) should be used, since the only T-DNA locus present in this genotype is situated within the gene of interest. Other R_1 genotypes would contain the second T-DNA locus that is presumed, like most other T-DNA insertions, not to cause an observable insertional mutation.

It is of interest to note that Potrykus and co-workers (Potrykus et al., 1985a) observed, after the direct transformation of protoplasts and their regeneration into mature tobacco plants, that two unexpected kanamycin-resistance segregation ratios occurred in the R_1 plants. These authors explained these results by the

A. Pollen lethal

		--	KK'	K-	-K'	♂
♀	--	S	X	R	X	1:1
		--	KK'	K-	-K'	♀
♂	--	S	R	R	R	3:1

B. Ova lethal

		--	KK'	K-	-K'	♂
♀	--	S	R	R	R	3:1
		--	KK'	K-	-K'	♀
♂	--	S	X	R	X	1:1

Table 9. Predicted phenotypes resulting from crossing TSS3-R_1 progeny as either pollen donors or pollen acceptors to wild-type plants

Part A of the table indicates the expected ratios if the mutation is pollen-lethal. The four possible gamete genotypes obtained from R_1 genotypes 2 and 8 (see Table 8) are shown horizontally. The symbol at the end of the row indicates whether the gametes are male or female. The gametes from the wild-type plant are indicated vertically; the sex of the gametes are indicated. The phenotypes of the progeny are indicated by either kanamycin-sensitive (S) or kanamycin-resistant (R). X indicates the absence of that class of progeny. The expected ratio of kanamycin-resistant to kanamycin-sensitive progeny is indicated. In part B of the Table, the phenotypes are calculated assuming that the mutation is ova-lethal. The same symbols are used as in part A. From these observation the gamete affected can be identified.

possible instability of transferred DNA. The method of direct transformation leads to the integration of foreign DNA at only one site in each transformed cell, giving rise to only one independently-segregating locus (Hain et al., 1985; Potrykus et al., 1985a). If the explanation proposed to interpret the results obtained in this work is applied to the non-Mendelian ratios obtained by Potrykus et al. (1985a) a similar conclusion can be reached. Potrykus et al. (1985a) failed to observe that the unusual ratios of 134:86 and 47:64 are statistically consistent with 2:1 and 1:1 ratios respectively. A plant with an insertion of foreign DNA into a gene required for either pollen or ova viability will give rise to a 1:1 segregation ratio of resistant to sensitive phenotypes in the R_1 generation when only one kanamycin-resistance locus is present. If however the inactivated gene is required for embryo viability, then embryos homozygous for the insertional mutation will not develop. Further, if the 0:4 resistant to sensitive ratios observed by Potrykus et al. (1985a) are considered in the same light, it is possible that some of these may be due to the insertional inactivation of genes required for both pollen and ova viability or required in two functional copies for embryo viability.

Similar observations have been made with transgenic mice. In two separate cases the disruption of genes essential for embryonic development by foreign DNA led to embryonic mortality and the absence of homozygous offspring in unusually small litters (Jaenisch et al., 1983; Wagner et al., 1983). In a third case, the foreign DNA was transmitted by females only. The sperm from males contained the inserted

DNA but did not transmit it to the next generation. A mouse embryonic-lethal mutation caused by insertional inactivation of the α1-collagen gene by a viral insertion event has also been described (Schnieke et al., 1983).

In plants insertional mutations due to transposable elements have been described (Federoff et al., 1984; O'Reilly et al., 1985), as have mutations produced by a procedure that only identifies insertion of a specially modified T-DNA into functional genes (Andre et al., 1986). There are no reports of normal T-DNA transfer giving rise to insertion mutants. One possible reason for this lack is the presence of large amounts of repetitive DNA in the plant genome. A second reason is the presence of multi-gene families in plant genomes, so that inactivation of one family member is unlikely to lead to a dramatic phenotypic change. A third possibility is the presence of two copies of each allele in the diploid plant. If an insertion event occurs at one locus essential only to haploid tissue, the cell may remain viable and regenerate into a mature plant; self-pollination of such a plant would only lead to progeny without the inserted DNA. These progeny would all be of the kanamycin-sensitive phenotype if only one insertion had occurred. If two insertion events had occurred, of which only one caused an insertion-mutation, the expected segregation of 1:15 would be reduced to 1:3; the plant would be scored as normal and presumed to have received only one T-DNA sequence. Some insertional mutations may therefore have been overlooked. If a single insertion event occurred in a gene that was essential in the homozygous form for cell viability, this cell would not regenerate into a whole

plant; again a mutational insertion event would not have been observed. It is therefore not surprising that no T-DNA insertion mutations have been described, other than those identified in haploid plants of Nicotiana plumbaginifolia (Andre et al., 1986), since it is probably not only a rare event, but recessive lethals are obscured by the dominant locus, and dominant lethals are unable to regenerate from tissue culture. The possibility that the disturbed Mendelian ratios observed in the progeny of TSS3 are due to the expression of a gene in haploid tissue, where the effect is not obscured by the wild-type allele, is therefore a novel and exciting observation.

One further consequence of this observation is that the inactivated gene is not likely to be a member of a multi-gene family since the other family members would obscure the observation. It should however be noted that N.tabacum is an allotetraploid, derived from a cross between N.sylvestris, and a member of the section Tomentosea, probably N.otophora (Goodspeed, 1954). The hybrid underwent spontaneous chromosome doubling to become tetraploid. The gametes therefore contain chromosomes from both parent plants and are not true haploids. The theory presented here relies on the non-complementation of the postulated mutation by the corresponding gene from the other parent plant species.

The genetic analyses that will resolve the problem of non-Mendelian segregation in the R_1 generation from TSS3 are currently underway. If the theory that an insertional mutation has occurred in TSS3 is correct, then the site of expression of the gene, pollen or ova, will also be

identified. Data from southern blots would also help resolve the question. The long-term work involved in the isolation of the gene may reveal that it encodes no more than an a gamete-specific enzyme; it may however lead to the identification of a more interesting gene involved in gamete development. The genetic screening of R_1 generations for the presence of gamete- or embryo-lethal mutations by perturbation of the segregation ratios may prove to be a valuable technique for the isolation of sequences involved in early development that would otherwise remain undiscovered.

SECTION IV

CONCLUSIONS

1. CONCLUSIONS

The data presented in this thesis can be summarised under three separate headings; these headings and the major points arising from the data are as follows:-

1) The isolation of two nuclear-encoded light-regulated genes from Pisum sativum.

   i) The characterisation, by restriction analysis, DNA sequencing and S1 nuclease analysis, of the genomic equivalent of the cDNA SSU60, namely, SS47.

   ii) The characterisation of a previously uncharacterised pea AB gene (AB13) by restriction endonuclease digestion and its partial sequence analysis.

2) The analysis of the expression of the SS and AB multi-gene familes in Pisum sativum.

   i) The demonstration that the member of the pea small subunit multi-gene family SS47 accounts for 50% of all small subunit transcripts in 9-day-old light-grown pea leaves.

   ii) The demonstration, by 5'- and 3'-S1 nuclease analysis, that defined members of the small

subunit multi-gene family account for different amounts of total SS mRNA in pea, and the estimation of their contribution to the total amount.

iii) The demonstration that the gene SS47, and the other members of the SS multi-gene family in pea are expressed in a light-regulated and organ-specific manner.

iv) The conclusion that there appear to be two transcription termination points for SS47 in light-grown leaf tissue, as defined by S1 nuclease analysis of RNA isolated from both pea and transgenic tobacco plants. There also appears to be a longer S1-protected SS transcript present in dark-grown pea apices as compared to light-grown pea leaf tissue. These observations await confirmation since they could be due to aberrant S1 nuclease digestion.

v) The observation that AB13, or a highly homologous gene, is expressed in a light-regulated and organ-specific manner in pea plants, and that other AB genes in P sativum var. Feltham First either share the same 5'-untranslated and coding sequences as AB13 or, are highly diverged from this gene on the basis of S1 nuclease analysis.

3) The transfer of the two genes isolated from P.sativum to plants of N.tabacum.

i) The establishment that SS47 is authentically expressed in transgenic tobacco plants. The degree of this expression in certain plants can reach amounts equivalent to those found in pea plants.

ii) The observation that inheritance of genes introduced into plant cells does not always follow the normal Mendelian rules. From this observation, a possible method for the identification and isolation of genes required during gamete formation and embryogenesis has been proposed.

Note added in proof:

During the final stages of preparing this manuscript, results similar to some of those presented in this thesis were published by other researchers [Fluhr, Moses, Morelli, Coruzzi & Chua (1986) EMBO J. 5, 2063-2073]. These workers presented the sequence of a small subunit gene from P.sativum var. Progress No.9 (rbcS-3A) that is identical to SS47 with the sole exception of one less base in the first intron. These workers also demonstrated the differential expression of the SS multi-gene family in pea by the use of 3'-S1 nuclease analysis, and the expression of rbcS-3A in transgenic petunia plants. The data presented here are in strong agreement with the results obtained by these authors.

REFERENCES

Abel, P. R., Nelson, R. S., De, B., Hoffmann, N., Rogers, S. G., Fraley R. T. & Beachy, R. N. (1986) Science 232, 738-743.

Akiyoshi, D. E., Klee, H., Amasino, R. M., Nester, E. W. & Gordon, M. P. (1984) Proc. Natl. Acad. Sci. USA 81, 5994-5998.

Allen, J.F. (1983) Trends Biochem.Sci. 8, 369-373.

Alt-Moerbe, J., Rak, B. & Schroder, J. (1986) EMBO J. 5, 1129-1135.

Amasino, R. M., Powell, A. L. T. & Gordon, M. P. (1984) Mol. Gen. Genet. 197, 437-446.

Amersham M13 Cloning and Sequencing Handbook (1984), Amersham UK.

Amersham Membrane Transfer and Detection Methods (1984), Amersham UK.

An, G., Watson, B. D., Stachel, S., Gordon, M. P. & Nester, E. W. (1985) EMBO J. 4, 277-284.

An, G., Ebert, P.R., Yi, B.-Y. & Choi, C.-H. (1986) Mol. Gen. Genet. 203, 245-250.

Andre, D., Colau, D., Schell, J., Van Montagu, M. & Hernalsteens, J.-P. (1986) Mol. Gen. Genet. 204, 512-518.

Ansorge, W. & De Mayer, L. (1980) J. Chromatog. 202, 45-53.

Apel, K. (1979) Eur. J. Biochem. 79, 183-188.

Apel, K., Kloppsteck, K. (1978) Eur. J. Biochem. 85, 581-588.

Balazs, E., Bouzoubaa, S., Guilley, H., Jonard, G., Paskowski, J. & Richards, K. (1985) Gene 40, 343-348.

Barta, A., Sanmergruber, K., Thompson, D., Hartmuth, K., Matzke, M. A. & Matzke, A. J. M. (1986) Plant Mol. Biol. 6, 347-357.

Barton, K. A., Binns, A. N., Matzke, A. J. M. & Chilton, M.-D. (1983) Cell 32, 1033-1043.

Batschauer, A., Mosinger, K., Kreuz, K., Dorr, I. & Apel, K. (1986) Eur. J. Biochem. 154, 625-634.

Beachy, R. N., Chen, Z.-L., Horsch, R. B., Rogers, S. G., Hoffmann, N. J. & Fraley, R. T. (1985) EMBO J. 4, 3047-3053.

Beaty, J. S., Powell, G. R., Lica, L., Regier, D. A., MacDonald, E. M. S., Hommes, N. G. & Morris, R. O. (1986) Mol. Gen. Genet. 203, 274-280.

Bedbrook, J. R., Smith, S. M. & Ellis, R. J. (1980) Nature 287, 692-697.

Bennett, J. (1981) Eur. J. Biochem. 118, 61-70.

Bennett, J. (1983) Biochem. J. 212, 1-13.

Bennett, J., Jenkins, G. I. & Hartley, M. R. (1984) J. Cell Biol. 25, 1-13.

Benoist, C., O'Hare, K., Breathnach, R., Chambon, P. (1980) Nucleic Acids Res. 8, 127-142.

Benton, W. D. & Davis, R. W. (1977) Science 196, 180-184.

Berry-Lowe, S. L., McKnight, T. D., Shah, D. M. & Meagher, R. B. (1982) J. Mol. Appl. Genet. 1, 483-498.

Bevan, M. W. (1984) Nucleic Acids Res. 12, 8711-8721.

Bevan, M. W. & Chilton, M.-D. (1982) Ann. Rev. Genet. 16, 357-384.

Bevan, M. W., Flavell, R. B. & Chilton, M.-D. (1983) Nature 304, 184-187.

Bevan, M. W., Mason, S. E. & Goetlet, P. (1985) EMBO J. 4, 1921-1926.

Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.

Binns, A. N., Sciaky, D. & Wood, H. N. (1982) Cell 31, 605-612.

Birnboim, H. C. & Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523.

Black, C. R., Kuleck, G. A. & Binns, A. N. (1986) Plant Physiol. 80, 145-151.

Bolton, G. W., Nester, E. W. & Gordon, M. P. (1986) Science 232, 983-985.

Bomhoff, G., Klapwijk, P. M., Kester, H. C. M., Schilperoort, R. A., Hernalsteens, J. P. & Schell, J. (1976) Mol. Gen. Genet. 145, 177-181.

Boyer, H. W. & Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.

Breathnach, R. & Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.

Brisson, N., Paszkowski, J., Penswick, J. R., Gronenborn, B., Potrykus, I. & Hohn, T. (1984) Nature 310, 511-514.

Broglie, R., Bellemare, G., Bartlett, S. G., Chua, N.-H. & Cashmore, A. R. (1981) Proc. Natl. Acad. Sci. USA 78, 7304-7308.

Broglie, R., Coruzzi, G., Lamppa, G., Kieth, B. & Chua, N.-H. (1983) Bio/technology 1, 55-61.

Broglie, R., Corruzi, G., Fraley, R. T., Rogers, S. G., Horsch, R. B., Niedermeyer, J. G., Finck, C. L., Flick, J. S. & Chua, N.-H. (1984) Science 224, 838-843.

Buchholz, W. G. & Thomashow, M. F. (1984a) J. Bacteriol. 160, 319-326.

Buchholz, W. G. & Thomashow, M. F. (1984b) J. Bacteriol. 160, 327-332.

Buchmann, I., Marner, F.-J., Schroder, G., Waffenschmidt, S. & Schroder, J. (1985) EMBO J. 4, 853-859.

Burke, J. F. (1984) Gene 30, 63-68.

Casadaban, M. J. & Cohen, S. N. (1984) Gene 30, 63-68.

Cashmore, A. R. (1983) in "Genetic Engineering of Plants- An Agricultural Perspective" (eds. Kosuge, T., Meredith, C. P. & Hollaender, A.) Basic Life Sciences vol. 26 Plenum Press N. Y., London pp 29-38.

Cashmore, A. R. (1984) Proc. Natl. Acad. Sci. USA 81, 2960-2964.

Chang, A. Y. C. & Cohen, S. N. (1978) J. Bacteriol. 134, 1141-1156.

Chilton, M.-D., Drummond, M. H., Merlo, D. J., Sciaky, D., Montoya, A. L., Gordon, M. P. & Nester, E. W. (1977) Cell 11, 263-271.

Chilton, M.-D., Montoya, A. L., Merco, D. J., Gordon, M. P. & Nester, E. W. (1978a) Plasmid 1, 254-269.

Chilton, M.-D., Drummond, M. H., Merlo, D. J. & Sciaky, D. (1978b) Nature 275, 147-149.

Chilton, M.-D., Saiki, R. K., Yadav, N., Gordon, M. P. & Quetier, F. (1980) Proc. Natl. Acad. Sci. USA 77, 4060-4064.

Chilton, W. S., Tempe, J., Matzke, M. & Chilton, M.-D. (1984) J. Bacteriol. 157, 357-362.

Chyi, Y.-S., Jorgensen, R. A., Goldstein, D., Tanksley, S. D. & Loaiza-Figueroa, F. (1986) Mol. Gen. Genet. 204, 64-69.

Comai, L., Facciotti, D., Hiatt, W. R., Thompson, G., Rose, R. E. & Stalker, D. M. (1985) Nature 317, 741-744.

Coruzzi, G., Broglie, R., Cashmore, A. & Chua, N.-H. (9183) J. Biol. Chem. 258, 1399-1402.

Coruzzi, G., Broglie, R., Edwards, C. & Chua, N.-H. (1984) EMBO J. 3, 1671-1679.

Czako, M. & Marton, L. (1986) Plant Mol. Biol. 6, 101-109.

Das, A., Stachel, S., Ebert, P., Allenza, P., Montoya, A. & Nester, E. (1986) Nucleic Acids Res. 14, 1355-1364.

Dean, C., van den Elzen, P., Tamaki, S., Dunsmuir, P. & Bedbrook, J. (1985a) EMBO J. 4, 3055-3061.

Dean, C., van den Elzen, P., Tamaki, S., Dunsmuir, P. & Bedbrook, J. (1985b) Proc. Natl. Acad. Sci. USA 82, 4964-4968.

Dean, C., van den Elzen, P., Dunsmuir, P. & Bedbrook, J. (1985c) in "Molecular Form and Function of the Plant Genome" (eds. van Vloten-Doting, C., Groot, G. S. P. & Hall, T. C.) Plenum Press. pp 167-174.

De Beuckeleer, M., Lemmers., De Vos, G., Willmitzer, L., Van Montagu, M. & Schell, J. (1981) Mol. Gen. Genet. 183, 283-288.

Deblaere, R., Bytebier, B., De Greve, H., Deboeck, F., Schell, J., Van Montagu, M. & Leemans, J. (1985) Nucleic Acids Res. 13, 4777-4788.

De Block, M., Herrera-Estrella, L., Van Montagu, M., Schell, J. & Zambryski, P. (1984) EMBO J. 3, 1681-1689.

De Block, M., Schell, J. & Van Montagu, M. (1985) EMBO J. 4, 1367-1372.

de Frammond, A. J. & Barton, K. A., Chilton, M. D. (1983) Bio/technology 1, 262-269.

de Frammond, A. J., Back, E. W., Chilton, W. S., Kayes, L. & Chilton, M.-D. (1986) Mol. Gen. Genet. 202, 125-131.

De Greve, H., Leemans, J., Hernalsteens, J. P., Thia-Toong, L., De Beucheleer, M., Willmitzer, L., Otten, L., Van Montagu, M. & Schell, J. (1982) Nature 300, 752-755.

De Greve, H., Dhaese, P., Seurinck, J., Lemmers, M., Van Montagu, M. & Schell, J. (1983) J. Mol.Appl. Genet. 1, 499-512.

Dellaporta, S. L., Wood, J. & Hicks, J. B. (1983) Plant Mol. Biol. rep. 1, 18-21.

Depicker, A., Van Montagu, M. & Schell, J. (1978) Nature 275, 150-153.

Depicker, A., De Wilde, M., De Vos, G., Van Montagu, M. & Schell, J. (1980) Plasmid 3, 193-211.

Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. & Goodman, H. M. (1982) J. Mol. Appl. Genet. 1, 561-574.

Depicker, A., Herman, L., Jacobs, A., Schell, J. & Van Montagu, M. (1985) Mol. Gen. Genet. 201. 477-484.

Deshayes, A., Herrera-Estrella, L. & Caboche, M. (1985) EMBO J. 4, 2731-2737.

De Vos, G., De Beucheleer, M., Van Montagu, M. & Schell, J. (1981) Plasmid 6, 149-253.

Dhaese, P., De Greve, H., Gielen, J., Seurinck, J., Van Montagu, M. & Schell, J. (1983) EMBO J. 2, 419-426.

Ditta, G., Stanfield, S., Corbin, D. & Helinski, D. R. (1980) Proc. Natl. Acad. Sci. USA 77, 7347-7351.

Dodgeson, J. B. & Wells, R. D. (1977) Biochemistry 16. 2374-2379.

Douglas, C. J., Staneloni, R. J., Rubin, R. A. & Nester, E. W. (1985) J. Bacteriol. 161, 850-860.

Dretzen, G., Bellard, M., Sassone-Corsi, P. & Chambon, P. (1981) Anal. Biochem. 112. 295-298.

Drummond, M. H. & Chilton, M.-D. (1978) J.Bacteriol. 136. 1178-1183.

Dunsmuir, P. (1985) Nucleic Acids Res. 13, 2503-2518.

Dunsmuir, P., Smith, S. M. & Bedbrook, J. (1983a) Nucleic Acids Res. 11, 4177-4183.

Dunsmuir, P., Smith, S. M. & Bedbrook, J. (1983b) J. Mol. Appl. Genet. 2, 285-295.

Ellis, J. G. & Murphy, P. J. (1981) Mol. Gen. Genet. 181. 36-43.

Ellis, R. J. (1979) Trends Biochem. Sci. 4, 241-244.

Ellis, R. J. (1981) Ann. Rev. Plant Physiol. 32, 111-137.

Ellis, R. J. (1983) Sub-cell. Biochem. 9 237-261.

Ellis, R. J. & Gatenby, A. A. (1984) in "The Genetic Manipulation of Plants and its Application to Agriculture", (eds. Lea, P. J. & Stewart, G. R.) Ann. Proc. Phytochem. Soc. Eur. vol. 23 Clarendon Press Oxford pp 41-60.

Engler, G., Depicker, A., Maenhaut, R., Villarroel-Mandiola, R., Van Montagu, M. & Schell, J. (1981) J. Mol. Biol. 152. 183-208.

Errede, B., Company, M., Ferchak, J. D., Huchinson III, C. A. & Yarnel, W. S. (1985) Proc. Natl. Acad. Sci. USA 82, 5423-5428.

Etkin, L. D., Pearman, B., Roberts, M. & Bektesh, S. L. (1984) Differentiation 26, 194-202.

Facciotti, D., O7Niel, J. K., Lee, S. & Shewmaker, C. K. (1985) Bio/Technology 3, 241-246.

Firmin, J. L. & Fenwick, G. R. (1978) Nature 276, 842-844.

Fluhr, R. & Chua N.-H. (1986) Proc. Natl. Acad. Sci. USA 83, 7358-7362.

Fluhr, R., Kuhlemeier, C., Nagy, F. & Chua, N.-H. (1986) Science 232. 1106-1112.

Fraley, R. T., Rogers, S. G., Horsch, R. B., Sanders, P. R., Flick, J. S., Adams, S. P., Bittner, M. L., Brand, L. A., Fink, J. S., Fry, J. S., Galluppi, G. R., Goldberg, S. B., Hoffman, N. L. & Woo, S. C. (1983) Proc. Natl. Acad. Sci. USA 80, 4803-4807.

Fraley, R. T., Horsch, R. B., Matzke, A., Chilton, M.-D., Chilton, W. S. & Sanders, P. R. (1984) Plant Mol. Biol. 3, 371-378.

Fraley, R. T., Rogers, S. G., Horsch, R. A., Eichholz, D. A., Flick, J. S.,

Fink, C. L., Hoffman, N. L. & Sanders, P. R. (1985) Bio/technology 3, 629-637.

French, R., Janda, M. & Ahlquist, P. (1986) Science 231. 1294-1297.

Frischauf, A.-M. Lehrach, H., Poustka, A. & Murray, N. (1983) J. Mol. Biol. 170, 827-842.

Fromm, M. E., Taylor, L. P. & Walbot, V. (1986) Nature 319. 791-793.

Funckes-Shippy, C. L. & Levine, A. D. (1985) in "Molecular Biology of the Photosynthetic Apparatus" (eds. Steinback, K. E., Bonitz, S., Arntzen, C. J. & Bogorad, L.) Cold Spring Harbour Lab. N. Y. pp 407-411.

Gallagher, T. f. & Ellis, R. J. (1982) EMBO J. 1, 149301498.

Gallagher, T. F., Jemkins, G. I. & Ellis, R. J. (1985) FEBS lett. 186, 241-245.

Gardner, R. C. & Knauf, V. C. (1986) Science 231, 725-727.

Garfinkel, D. J. & Nester, E. W. (1980) J. Bacteriol. 144, 732-743.

Garfinkel, D. J., Simpson, R. B., Ream, L. W., White, F. F., Gordon, M. P. & Nester, E. W. (1981) Cell 27, 143-153.

Gelvin, S. B., Thomashow, M. F., McPherson, J. C., Gordon, M. P. & Nester, E. W. (1982) Proc. Natl. Acad. Sci. USA 79, 76-80.

Gelvin, S. B., Karcher, S. J. & DiRiti, V. (1983) Nucleic Acids Res. 11, 159-174.

Gielen, J., De Beuckeleer, M., Seurinck, J., Deboeck, F., De Greve, H., Lemmers, M., Van Montagu, M. & Schell, J. (1984) EMBO J. 3, 835-846.

Golden, S. S. & Haselkorn, R. (1985) Science 299, 1104-1109.

Goldsborough, P. B., Gelvin, S. B. & Larkin, B. A. (1986) Mol. Gen. Genet. 202, 374-381.

Gollmer, I. & Apel, K. (1983) Eur. J. Biochem. 133, 309-313.

Goodman, T.H. (1954) The Genus Nicotiana. Waltham, Mass.: Chronica. botanica.

Gronenborn, B., Gardner, R. C., Schaefer, S. & Sheperd, R. J. (1981) Nature 294, 773-776.

Grunstein, M. & Hogness, D. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965.

Hain, R., Sabel, P., Czernilofsky, A. P., Steinbiss, H. H., Herrera-Estrella, L. & Schell, J. (1985) Mol. Gen. Genet. 199, 161-168.

Hanahan, D. (1983) J. Mol. Biol. 166, 557-580.

Hernalsteens, J.-P., Thia-Toong, L., Schell, J. & Van Montagu, M. (1984) EMBO J. 3, 3039-3041.

Herrera-Estrella, L., Depicker, A., Van Montagu, M. & Schell, J. (1983a) Nature 303, 209-213.

Herrera-Estrella, L., De Block, M., Messens, E., Hernalsteens, J.-P., Van Montagu, M. & Schell, J. (1983b) EMBO J. 2, 987-995.

Herrera-Estrella, L., Van den Broeck, G., Maenhaut, R., Van Montagu, M. & Schell, J. (1984) Nature 310, 115-120.

Highfield, P. E. & Ellis, R. J. (1978) Nature 271, 420-424.

Hille, J., Verheggen, F., Roelvink, P., Frassen, H., van Kammen, A. & Zabel, P. (1986) Plant Mol. Biol. 7, 171-176.

Hoekema, A., Hirsch, P. R., Hooykaas, P. J. J. & Schilperoort, R. A. (1983) Nature 303, 179-180.

Hoekema, A., de Pater, B. S., Fellinger A. J., Hooykaas, P. J. J. & Schilperoort, R. A. (1984a) EMBO J. 3, 3034-3047.

Hoekema, A., Roelvink, P. W., Hooykaas, P. J. J. & Schilperoort, R. A. (1984b) EMBO J. 3 2485-2490.

Hoekema, A., Hooykaas, P. J. J. & Schilperoort, R. A. (1984c) J. Bacteriol. 158, 383-385.

Hoekema, A., van Haaren, M. J. J., Fellinger, A. J., Hooykaas, P. J. J. & Schilperoort, R. A. (1985) Plant Mol Biol. 5, 85-89.

Holsters, M., Silva, B., Van Vliet, F., Genetello, C., De Block, M., Dhaese, P., Depicker, A., Inze, D., Engler, G., Villarroel, R., Van Montagu, M. & Schell, J. (1980) Plasmid 3, 212-230.

Holsters, M., Villarroel, R., Gielen, J., Seurinck, J., De Greve, H., Van Montagu, M. & Schell, J. (1983) Mol. Gen. Genet. 190, 35-40.

Hooykaas, P. J. J., Hofker, M., den Dulk-Ras, H. & Schilperoort, R. A. (1984) Plasmid 11, 195-205.

Hooykaas-van Slogteren, G. M. S., Hooykaas, P. J. J. & Schilperoort, R. A. (1984) Nature 311, 763-764.

Horsch, R. B., Fraley, R. T., Rogers, S. G., Sanders, P. S., Lloyd, A. &

Hoffmann, N. (1984) Science 233, 496-498.

Horsch, R. B., Fry, J. E., Hoffmann, N., Eichholz, D., Rogers, S. G. & Fraley, R. T. (1985) Science 227, 1229-1231.

Hull, R. & Davies, J. W. (1983) Adv. Virus Res. 28, 1-33.

Inze, D., Follon, A., Van Lijsebettens, M., Simoens, C., Genetello, O., Van Montagu, M. & Schell, J. (1984) Mol. Gen. Genet. 194, 265-274.

Iyer, V. N., Klee, H. J. & Nester, E. W., (1982) Mol. Gen. Genet. 188, 418-424.

Jaenisch, R., Harbers, K., Schneiker, A., Lohler, J., Chumakov, I., Jahner, D., Grotkopp, D. & Hoffman, E. (1983) Cell 32, 209-216.

Jen, G. C. & Chilton, M.-D. (1986a) J. Bacteriol. 166. 491-499.

Jen, G. C. & Chilton, M.-D. (1986b) Proc. Natl. Acad. Sci. USA 83, 3895-3899.

Jenkins, G. I., Hartley, M. R. & Bennett, J. (1983) Philos. Trans. R. Soc. London Ser. B 303, 419-431.

Jones, J. D. G., Dunsmuir, P. & Bedbrook, J. (1985) EMBO J. 4, 2411-2418.

Joos, H., Inze, D., Caplan, A., Sormann, M., Van Montagu, M. & Schell, J. (1983) Cell 32, 1057-1067.

Karcher, S. J., DiRita, V. J. & Gelvin, S. B. (1984) Mol. Gen. Genet. 194. 159-165.

Karlin-Neuman, G. A., Kohorn, B. D., Thornber, J. P. & Tobin, E. M. (1985) J. Mol. Appl. Genet. 3, 45-61.

Kato, S., Anderson, R. A. & Camerini-Otero, R. D. (1986) Mol. Cell Biol. 6, 1787-1795.

Kaufman, L. S., Thompson, W. F. & Briggs, W. R. (1984) Science 226, 1147-1149.

Kaufman, L. S., Watson, J. C., Briggs, W. R. & Thompson, W. F. (1985) in "Molecular Biology of the Photosynthetic Apparatus (eds. Steinbeck, K. E., Bonitz, S., Arntzen, C. J. & Bogorad, L.) Cold Spring Harbour Lab. N. Y. pp 367-372.

Kawashima, N. & Wildman, S. G. (1972) Biochim. Biophys. Acta 262. 42-49.

Kerr, A., (1969) Nature 223, 1175-1176.

Kerr, A., (1971) Physiol. Plant Pathol. 1, 141-146.

Klee, H. J., Gordon, M. P. & Nester, E. W. (1982) J. Bacteriol. 150. 327-331.

Klee, H. J., White, F. F., Iyer, V. N., Gordon, M. P. & Nester, E. W. (1983) J. Bacteriol. 153. 878-883.

Klee, H. J., Yanofsky, M. F. & Nester, E. W. (1985) Bio/technology 3, 637-642.

Knauf, V., Yanofsky, M., Montoya, A. & Nester, E. W. (1984) J. Bacteriol. 160, 564-568.

Koncz, C., De Greve, H., Andre, D., Deboeck, F., Van Montagu, M. & Schell, J. (1983) EMBO J. 2, 1597-1603.

Koncz, C., Kreuzaler, F., Kalman, Z. & Schell, J. (1984) EMBO J. 3, 1029-1037.

Koncz, C. & Schell, J. (1986) Mol. Gen. Genet. 204, 383-396.

Koukolikova-Nicola, Z., Shillito, R. D., Hohn, B., Wang, K., Van Montagu, M. & Zambryski, P. (1985) Nature 313, 191-106.

Kozak, M. (1981) Nucleic acids Res. 9, 5233-5252.

Koziel, M. G., Adams, T. L., Hazlet, M. A., Damm, D., Miller, J., Dahlbeck, D., Jayne, S. & Staskawicz, B. (1984) J. Mol. Appl. Genet. 2, 549-562.

Krens, F. A., Molendijk, L., Wullems, G. J. & Schilperoort, R. A. (1982) Nature 296, 179-180.

Kung, S.-D., Thornber, J. P. & Wildman, S. G. (1972) FEBS Lett. 24, 185-188.

Lamb, F. I. (1984) PhD Thesis, University of Warwick.

Lamppa, G. K., Morelli, G. & Chua, N.-H. (1985a) Mol. Cell Biol. 5, 1370-1378.

Lamppa, G. K., Nagy,F. & Chua, N.-H. (1985b) Nature 316, 750-752.

Langridge, P. (1981) FEBS Lett. 123, 85-89.

Leemans, J., Debleare, R., Willmitzer, L., De Greve, H., Hernalsteens, J. P., Van Montagu, M. & Schell, J. (1982) EMBO J. 1, 147-152.

Lemmers, M., De Beuckelerr, M., Holsters, M., Zambryski, P., Depicker, A., Hernalsteens, J. P., Van Montagu, M. & Schell, J. (1980) J. Mol. Biol. 144, 353-376.

Leutwiler, L. S., Meyerowitz, E. M. & Tobin, E. M. (1986) Nucleic Acids Res. 14, 4051-4064.

Lichtenstein, C. P. & Draper, J. (1981) in "DNA Cloning Volume II, A Practical Approach" (ed. D. M. Glover) IRL press pp 67-119.

Lorz, H., Baker, B. & Schell, J. (1985) Mol. Gen. Genet. 199. 178-182.

Lui, S.-T., Perry, K. L., Schardl, C. L. & Kado, C. I. (1982) Proc. Natl. Acad. Sci. USA 79, 2812-2816.

Lundquist, R. C., Close, T. J. & Kado, C. I. (1984) Mol. Gen. Genet. 193, 1-7.

Machida, Y., Usami, S., Yamamoto, A., Niwa, Y. & Takebe, I. (1986) Mol. Gen. Genet. 204, 374-382.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual. CSH Lab.

Marton, L., Wullems, G. J., Molendijk, L. & Schilperoort, R. A. (1979) Nature 277, 129-131.

Matzke, M., Susani, M., Binns, A., Lewis, E., Rubenstein, I. & Matzke, A. (1984) EMDO J. 3, 1525-1531.

Mazur, B. J. & Chui, C.-F. (1985) Nucleic Acids Res. 13. 2373-2386.

McKnight, G. S. & Palmiter, R. D. (1979) J. Biol. Chem. 254, 9050-9058.

Merlo, C. J., Nutter, R. C., Montoya, A. L., Garfinkel, D. J., Drummond, M. H., Chilton, M.-D., Gordon, M. P. & Nester, E. W (1980) Mol. Gen. Genet. 177, 637-643.

Messens, E., Lenearts, A., Van Montagu, M. & Hedges, R. W. (1985) Mol. Gen. Genet. 199, 344-348.

Messing, J., Crea, R. & Seeburg, P. H. (1981) Nucleic Acids Res. 9, 309-321.

Messing, J., Geraghty, D., Heidecker, G., Hu, N.-T., Kridl, J. & Rubenstein, I. (1983) in "The Genetic Engineering of Plants; an agricultural perspective" (eds. Kosuge, T., Meredith, C. & Holleander, A.) Plenum N.Y. 211-227.

Miziorko, H. M. & Lorimer, G. H. (1983) Ann. Rev. Biochem. 52, 507-535.

Montoya, A. L., Chilton, M.-D., Gordon, M. P., Sciaky, D. & Nester, E. W. (1977) J.Bacteriol. 129, 101-107.

Moore, L., Warren, G. & Strobel, G. (1979) Plasmid 2, 617-626.

Morelli, G., Nagy, F., Fraley, R. T., Rogers, S. G. & Chua, N.-H. (1985) nature 315, 200-204.

Mount, S. M. (1982) Nucleic Acids Res. 10, 459-472.

Murai, N., Sutton, D. W., Murray, M. G., Slightom, J. L., Merlo, D. J., Reichrt, N. A., Sengupta-Gopalan, C., Stock, C. A., Barker, R. F., Kemp, J. D. & Hall, T. C. (1983) Science 222, 476-481.

Musgrove, J. E. & Ellis, R. J. (1986) Philos. Trans. R. Soc. London 313, 419-428.

Nagy, F., Morreli, G., Fraley, R. T., Rogers, S. G. & Chua, N.-H. (1985) EMBO J. 4, 3063-3068.

Nagy, F., Kay, S. A., Boutry, M., Hsu, M.-Y. & Chua, N.-H. (1986) EMBO J. 5, 1119-1124.

Nester, E. W., Gordon, M. P., Amasino, R. M. & Yanofsky, M. F. (1984) Ann. Rev. Plant. Physiol. 35, 387-413.

Nierzwicki-Bauer, S., Curti, S. e. & Haselkorn, R. (1984) Proc. natl. acad. sci. USA 81, 5961-5965.

Ohta, Y. (1986) Proc. Natl. Acad. Sci. 83, 715-719.

Okker, R. J. H., Spaink, H., Hille, J., van Brussel, T. A. N., Lugtenberg, B. & Schilperoort, R. A. (1984) Nature 312, 564-566.

Ooms, G., Klapwijk, P. M., Poulis, J. A. & Schilperoort, R. A. (1980) J. Bacteriol. 144, 82-91.

Ooms, G., Hooykaas, P. J. J., Moolenaar, G. & Schilperoort, R. A. (1981) Gene 14, 33-50.

Ooms, G., Bakker, A., Molendijk, L., Wullems, G. J., Gordon, M. P., Nester, E. W. and Schilperoort, R. A. (1982) Cell 30, 589-597.

O'Reilly, C., Shepherd, N. S., Pereira, A., Schwarz-Sommer, Z., Bertram, I., Robertson, D. S., Peterson, P. A. & Saedler, H. (1985) EMBO J. 4, 877-882.

Otten, L., De Greve, H., Hernalsteens, J. P., Van Montagu, M., Schneider, O., Straub, J. & Schell, J. (1981) Mol. Gen. Genet. 183, 209-213.

Otten, L., Piotrowiak, G., Hooykaas, P., Dubois, M., Szegedi, E. & Schell, J. (1985) Mol. Gen. Genet. 199, 189-193.

Palmiter, R. D. & Brinster, R. L. (1985) Cell 41, 343-345.

Palmiter, R. D., Wilkie, T. M., Chen, H. Y. & Brinster, R. C. (1984) Cell 36, 869-877.

Paszkowski, J., Shillito, R. D., Saul, M., Mandak, V., Hohn, T., Hohn, B. and Potrykus, I. (1984) EMBO J. 3, 2717-2722.

Paszkowski, J., Pisan, B., Shillito, R., Hohn, T., Hohn, B. and Potrykus, I. (1986) Plant. Mol. Biol. 6, 303-312.

Peralta, E. G. & Ream, L. W. (1985) Proc. Natl. Acad. Sci. USA 82, 5112-5116.

Peralta, E. G., Hellmiss, R. & Ream, L. W. (1986) EMBO J. 5, 1137-1142.

Petit, A., Delahate, S., Tempe, J. & Morel, G. (1970) Physiol. Veg. 8, 205-213.

Petit, A., David, C., Dahl, G. A., Ellis, J. G., Guyon, P., Casse-Delbart, F. & Tempe, J. (1983) Mol. Gen. Genet. 190, 204-214.

Petit, A., Berkaloff, A. & Tempe, J. (1986) Mol. Gen. Genet. 202, 388-393.

Pichersky, E., Bernatzky, R., Tanksley, S. D. & Cashmore, A. R. (1986) Proc. Natl. Acad. Sci. USA 83, 3880-3884.

Pietrzak, M., Shillito, R. D., Hohn, T. & Potrykus, I. (1986) Nucleic Acids Res. 14, 5857-5868.

Pinck, L., Fleck, J., Pinck, M., Hadidane, R. & Hirth, L. (1983) FEBS Lett. 154, 145-148.

Polans, N. O., Weeden, N. F. & Thompson, W. F. (1985) Proc. Natl. Acad. Sci. USA 82, 5083-5087.

Potrykus, I., Paszkowski, J., Saul, M. W., Petruska, J. & Shillito, R. D. (1985a) Mol. Gen. Genet. 199, 169-177.

Potrykus, I., Saul, M. W., Petruska, J., Paszkowski, J. & Shillito, R. D. (1985b) Mol. Gen. Genet. 199, 183-188.

Pratt, L. J. (1982) Ann. Rev. Plant Physiol. 33, 557-574.

Robinson, C. & Ellis, R. J. (1984) Eur. J. Biochem. 142. 343-346.

Rigby, P. N. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) J. Mol. Biol. 113, 237-251.

Rusche, J. R. & Howard-Flanders, P. (1985) Nucleic Acids Res. 6, 1997-2008.

Sailland, A., Amiri, I. & Freyssint, G. (1986) Plant Mol. Biol. 7, 213-218.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Sasaki, Y., Ishiye, M., Sakahama, T. & Kamikubo, T. (1981) J. Biol. Chem. 256, 2315-2320.

Schafer, W., Weising, K. & Kahl, G. (1984) EMBO J. 3, 373-376.

Schmidt, G. W., Bartlett, S. G., Grossman, A. R., Cashmore, A. R. & Chua, N.-H. (1981) J. Cell. Biol. 91, 468-478.

Schneike, A., Harbers, K. & Jaenisch, R. (1983) Nature 304, 315-320.

Schoffl, F. & Baumann, G. (1985) EMBO J. 4, 1119-1124.

Schreier, P. H., Seftor, E. A., Schell, J. & Bohnert, H. J. (1985) EMBO J. 4, 25-32.

Schroder, J., Schroder, G., Huisman, H. & Schilperoort, R. A. (1981) FEBS lett. 129, 166-168.

Schroder, G., Klipp, W., Hillebrand, A., Ehring, R., Koncz, C. & Schroder, J. (1983) EMBO J. 2, 403-409.

Sengupta-Gopalan, C., Reichert, N. A., Barker, R. F., Hall, T. C. & Kemp, J. D. (1985) Proc. Natl. Acad. Sci. USA 82, 3320-3324.

Shah, D., Horsch, R., Klee, H., Kishmore, G., Winter, J., Turner,N., Hironaka, C., Sanders, P., Gasser, C., Aykent, S., Siegel, N., Rogers, S. & Fraley, R. (1986) Science 233, 478-481.

Shaw, C. H., Leemans, J., Shaw, C. H., Van Montagu, M. & Schell, J. (1983) Gene 23, 315-330.

Shaw, C. H., Watson, M. D., Carter, G. H. & Shaw, C. H. (1984) Nucleic Acids Res. 12, 6031-6041.

Shepard, J. F. (1983) Scientific American 248, 112-121.

Shinozaki, K. & Sugiura, M. (1983) Nucleic Acids Res. 11, 6957-6964.

Silverthorne, J. & Tobin, E. M. (1984) Proc. Natl. Acad. Sci. USA 81, 1112-1116.

Simpson, R. B., O'Hara, P. J., Kwok, W., Montoya, A. L., Lichtenstein, C., Gordon, M. P. & Nester, E. W. (1982) Cell 29, 1005-1014.

Simpson, J., Timko, M. P., Cashmore, A. R., Schell, J., Van Montagu, M. & Herrera-Estrella, L. (1985) EMBO J. 4, 2723-2729.

Simpson, R. B., Spielmann, A., Margossian, L. & McKnight, T. D. (1986a) Plant Mol. Biol. 6, 403-415.

Simpson, J., Van Montagu, M. & Herrera-Estrella, L. (1986b) Science 233, 34-38.

Smith, E. F. & Townsend, C. O. (1907) Science 25, 671-673.

Smith, S. M., Bedbrook, J. & Spiers, J. (1983) Nucleic Acids Res. 11, 8719-8734.

Smith, S. M. & Ellis, R. J. (1981) J. Mol. Appl. Genet. 1, 127-137.

Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.

Spena, A., Hain, R., Ziervogel, U., Saedler, H. & Schell, J. (1985) EMBO J. 4, 2739-2743.

Stachel, S. F., Messens, E., Van Montagu, M. & Zambryski, P. (1985) Nature 318, 624-629.

Stachel, S. F. & Nester, E. W. (1986) EMBO J. 5, 1445-1454.

Stachel, S. F., Nester, E. W. & Zambryski, P. (1986a) Proc. Natl. Acad. Sci. USA 83, 379-383.

Stachel, S. F., Timmerman, B. & Zambryski, P. (1986b) Nature 322, 706-712.

Stiekema, W. J., W'impee, C. F. & Tobin, E. M. (1983a) Nucleic Acids Res. 11, 8051-8061.

Stiekema, W. J., Wimpee, C. F., Silverthorne, J. & Tobin, E. M. (1983b) Plant Physiol. 72, 717-724.

Stougard-Jensen, J., Mardier, K. A., Otten, L. & Schell, J. (1986) Nature 321, 669-674.

Teeri, T. H., Herrera-Estrella, L., Depicker, A., Van Montagu, M. & Palva, E. T. (1986) EMBO J. 5, 1755-1760.

Tempe, J. & Petit, A. (1985) in "Genetic Engineering of Plants and Microorganisms Important for Agriculture" (eds. Magnien, E. & de Nettancourt, D.) Martinus Nijhoff pub. 82-82.

Thomashow, M. F., Nutter, R., Montoya, A. L., Gordon, M. P. & Nester, E. W. (1980a) Cell 19, 729-739.

Teyssendier de la Serve, B., Axelos, M. & Peaud-Leonel, C. (1985) Plant Mol. biol. 5, 155-163.

Thomashow, M. F., Nutter, R., Postle, K., Chilton, M.-D., Blattner, F. R., Powell, A., Gordon, M. P. & Nester, E. W. (1980b) Proc. Natl. Acad. Sci. USA 77, 6448-6452.

Thomashow, L. S., Reeves, S. & Thomashow, M. F. (1984) Proc. Natl. Acad. Sci. USA 81, 5071-5075.

Thomashow, M. F., Hugly, S., Buchholz, W. G. & Thomashow, L. S. (1986) Science 231, 616-618.

Thompson, W. F., Everett, M., Polans, N. O., Jorgensen, R. a. & Palmer, J. D. (1983) Planta 158, 487-500.

Thomson, J. A. (1986) Tibtech 4, 219-224.

Timko, M., Kausch, A., Hand, J., Cashmore, A., Herrera-Estrella, L., Van den Broeck, G. & Van Montagu, M. (1985a) in "Biosynthesis of the Photosynthetic Apparatus" (eds. Thornber, J., Staehelin, L. & Hallich, R.) Liss N.Y. pp 381-396.

Timko, M. P. & Cashmore, A. R. (1983) in "Plant Molecular Biology" (ed. Goldberg, R. B.) Liss N.Y. pp 403-412.

Timko, M. P., Kausch, A. P., Castresana, C., Fassler, J., Herrera-Estrella, L., Van den Broeck, G., Van Montagu, M., Schell, J. & Cashmore, A. R. (1985b) Nature 318, 579-582.

Tobin, E. M. (1978) Proc. Natl. Acad. sci. USA 75, 4749-4753.

Tobin, E. M. (1981) Plant Mol. Biol. 1, 35-51.

Tobin, E. M. & Suttie, J. L. (1980) Plant Physiol. 65, 641-647.

Tobin, E. M. & Silverthorne, J. (1985) Ann. Rev. Plant. Physiol. 36, 569-593.

Tobin, E. M., Wimpee, C. F., Silverthorne, J., Stiekema, W. J., Neuman, G. a. & Thornber, J. P. (1984) in "Biosynthesis of the Photosynthetic Apparatus" (eds. Thornber, J. P., Staehelin, L. A. & Hallich, R.) Liss N.Y. pp 325-334.

Tobin, E. M., Wimpee, C. F., Karlin-Neuman, G. A., Silverthorne, J. & Kohorn, B. D. (1985) in "Molecular Biology of the Photosynthetic Apparatus" (eds. Steinbeck, K. E., Bonitz, S., arntzen, C. J. & Bogorad, L.) Cold Spring Harbour Lab. N.Y. pp 373-380.

Townes, T. M., Lingrel, J. B., Chen, H. Y., Brinster, R. L. & Palmiter, R. D. (1985) EMBO J. 4, 1715-1723.

Twigg, A. J. & Sherratt, D. (1980) Nature 283, 216-218.

Ursic, D., Slightom, J. L. & Kemp, J. D. (1983) Mol. Gen. Genet. 190, 494-503.

Van den Broek, G., Timko, M. P., Kausch, A. P., Cashmore, A. R., Van Montagu, M. & Herrera-Estrella, L. (1985) Nature 313, 358-363.

van den Elzen, P. J. M., Lee, K. Y., Townsend, J. & Bedbrook, J. (1985a) Plant Mol. Biol. 5, 149-154.

van den Elzen, P. J. M., Townsend, J., Lee, K. Y. & Bedbrook, J. (1985b) Plant Mol. Biol. 5, 299-302.

Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M. & Schell, J. (1983) EMBO J. 2, 411-417.

Van Larebeke, N., Engler, G., Holsters, M., Van den Elsacker, S., Zaenen, I., Schilperoort, R. A. & Schell, J. (1974) Nature 252, 169-170.

Van Larebeke, N., Genetello, C., Schell, J., Schilperoort, R. A., Hermans, A. K., Hernalsteens, J. P. &

Van Montagu, M. (1975) Nature 255, 742-743.

Van Montagu, M. & Schell, J. (1982) Current Topics in Microbiology and Immunology 96, 237-254.

van Slogteren, G. M. S., Hooykaas, P. J. J. & Schilperoort, R. A. (1984a) Plant Mol. Biol. 3, 337-344.

van Slogteren, G. M. S., Hooykaas, P. J. J. & Schilperoort, R. A. (1984b) Plant Mol. Biol. 3, 333-336.

Velten, J. & Schell, J. (1985) Nucleic Acids res. 13, 6981-6998.

Virts, E. L. & Gelvin, S. B. (1985) J. Bacteriol. 162, 1030-1038.

Wagner, E. F., Covarrubias, L., Stewart, T. A. & Mintz, B. (1983) Cell 35, 647-655.

Waldron, C., Murphy, E. B., Roberts, J. L., Gustafson, G. D., Armour, S. L. & Malcolm, S. K. (1985) Plant Mol. Biol. 5, 103-108.

Wang, K., Herrera-Estrella, L., Van Montagu, M. & Zambryski, P. (1984) Cell 24, 455-462.

Watson, B., Currier, T. C., Gordon, M. P., Chilton M.-D. & Nester, E. W. (1975) J. Bacteriol. 123, 255-264.

Weiher, H., Konig, M. & Gruss, P. (1983) Science 219, 626-631.

Werr, W. & Lorst, H. (1986) Mol. Gen. Genet. 202, 471-475.

White, P. R. & Braun, A. C. (1941) Science 94, 239-241.

White, P. R. & Braun, A. C. (1942) Cancer Res. 2, 597-617.

Willmitzer, L., De Beuckeleer, M., Lemmers, M., Van Montagu, M. & Schell, J. (1980) Nature 287, 359-361.

Willmitzer, L., Schmalenbach, W. & Schell, J. (1981) Nucleic Acids Res. 9, 4801-4812.

Willmitzer, L., Simons, G. & Schell, J. (1982) EMBO J. 1, 139-146.

Willmitzer, L., Dhaese, P., Schreier, P., Schmalenbach, W., Van Montagu, M. & Schell, J. (1983) cell 32, 1045-1056.

Wimpee, C. F., Stiekema, W. J. & Tobin, E. M. (1983) in "Plant Molecular Biology" (ed. Goldberg, R. B.) Liss N.Y. pp 391-401.

Wood, W. B. (1966) J. Mol. Biol. 16, 118-133.

Wullems, G. J., Molendijk, L., Ooms, G. & Schilperoort, R. A. (1981a) Proc. Natl. Acad. Sci. USA 78, 4344-4348.

Wullems, G. J., Molendijk, L., Ooms, G. & Schilperoort, R. A. (1981b) Cell 24, 719-729.

Yadav, M. S., Vanderleyden, J., Bennett, D. R., Barnes, W. M. & Chilton M.-D. (1982) Proc. Natl. Acad. Sci. USA 79, 6322-6326.

Yang, F.-M. & Simpson, R. B. (1981) Proc. Natl. Acad. Sci. USA 78, 4151-4155.

Zambryski, P., Holsters, M., Kruger, K., Depicker, A. & Schell, J. (1980) Science 209, 1385-1391.

Zaenen, I., Van Larebeke, N., Teuchy, H., Van Montagu, M. & Schell, J. (1974) J. Mol. Biol. 86, 109-127.

Zimmermann, U. & Vienken, J. J. (1982) J. Membrane Biol. 67, 165-182.